DOCUMENT RESUME

ED 064 668 24 CG 007 518

AUTHOR	Resnick, Lauren B., Ed.
TITLE	Hierarchies in Children's Learning.
INSTITUTION	Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
BUREAU NO	BR-5-0253
PUB DATE	71
NOTE	119p.; Paper presented at the American Educational Research Association, Minneapolis, Minnesota, in March of 1970
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	Child Development; *Children; *Learning; Learning Activities; *Learning Characteristics; Learning Experience; Learning Plateaus; *Learning Processes; Learning Readiness; Learning Theories; Problem Solving

ABSTRACT

This monograph reports the proceedings of a symposium that considered issues in the use of learning hierarchies in both psychological and educational research. The opening paper presents a brief overview of research on learning hierarchies. Issues considered include the use of behavior analysis as a basis for generating hierarchies, the extent to which hierarchy theory can take into account individual differences in learning patterns, and the implications of cumulative learning of increasingly more complex tasks for theories of cognitive development. Different methods of validating hypothesized hierarchies are discussed. In the succeeding papers, four instances of research on hierarchies are presented, each representing a different approach to the problem of hypothesizing and testing hierarchical relations among cognitive tasks. Several discussants, representing the developmental, learning, and psychometric points of view, then consider implications of the research reported. (Author)

BR 50253
PA24
CG

LEARNING RESEARCH AND DEVELOPMENT CENTER

1972/10

HIERARCHIES IN CHILDREN'S LEARNING

LAUREN B. RESNICK
EDITOR

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

CG 007 518

ED 064668

ED 064668

HIERARCHIES IN CHILDREN'S LEARNING

Lauren B. Resnick

Editor

Learning Research and Development Center

University of Pittsburgh

1971

The papers and discussions presented herein constitute the proceedings of a symposium held at the 1970 meetings of the American Educational Research Association, Minneapolis, Minnesota. Preparation of the manuscript was supported by grants from the Ford Foundation and by the Learning Research and Development Center supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed do not necessarily reflect the position or policy of the Office of Education and no official endorsement should be inferred.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iii

ABSTRACT iv

ISSUES IN THE STUDY OF LEARNING HIERARCHIES
Lauren B. Resnick 1

PSYCHOMETRIC STUDIES IN THE VALIDATION OF AN EARLY LEARNING CURRICULUM
Margaret C. Wang 21

THE EFFECT OF SEQUENCE IN THE ACQUISITION OF THREE SET RELATIONS: AN EXPERIMENT WITH PRESCHOOLERS
A. Edward Uprichard 41

HIERARCHICAL RELATIONSHIPS AND TRANSFER IN LEARNING MULTIPLICATIVE CLASSIFICATION TASKS
Alexander W. Siegel and Esther Kresh 58

A STUDY OF SUBORDINATE SKILLS IN SCIENCE PROBLEM SOLVING
Virginia K. Wiegand 79

DISCUSSION
John B. Carroll 94

DISCUSSION
Robert Glaser 101

DISCUSSION
Millie Almy 109

ACKNOWLEDGEMENTS

The authors would like to thank Donna Rottman for preparing the figures and Cathlene Hardaway and Shellie Williams for typing the manuscript.

HIERARCHIES IN CHILDREN'S LEARNING

Lauren B. Resnick

Editor

University of Pittsburgh

ABSTRACT

This monograph reports the proceedings of a symposium that considered issues in the use of learning hierarchies in both psychological and educational research. The opening paper presents a brief overview of research on learning hierarchies. Issues considered include the use of behavior analysis as a basis for generating hierarchies, the extent to which hierarchy theory can take into account individual differences in learning patterns, and the implications of cumulative learning of increasingly more complex tasks for theories of cognitive development. Different methods of validating hypothesized hierarchies are discussed.

In the succeeding papers, four instances of research on hierarchies are presented, each representing a different approach to the problem of hypothesizing and testing hierarchical relations among cognitive tasks. Several discussants, representing the developmental, learning, and psychometric points of view, then consider implications of the research reported.

ISSUES IN THE STUDY OF LEARNING HIERARCHIES

Lauren B. Resnick

University of Pittsburgh

Since Gagné first used the term "hierarchy" in his theory of how human beings acquired complex skills and knowledge (Gagné, 1962), there has been a continuing increase in the application of hierarchy theory to problems of instruction and evaluation. Simultaneously, but apparently independently, developmental psychologists have begun to use concepts of hierarchical dependency in studying sequences of cognitive and psychosocial development. This symposium will examine some current research on learning hierarchies and consider the implications of this research for theories of cognitive development and instructional psychology.

Definition of a Hierarchy

Hierarchies have been differently defined by different investigators, in accord with their theoretical and applied interests. Learning psychologists and instructional designers tend to define hierarchies in terms of asymmetrical transfer relationships between two or more tasks. Thus, two tasks are considered to be hierarchically related if a) one task is easier to learn than the other, and b) learning the simpler task first produces positive transfer in learning the more complex task. For example, learning to count is demonstrably easier than learning to add. A child will also learn addition more quickly if he is competent in counting than if he does not yet know how to count; counting, in other words, provides positive transfer to learning addition. Counting and adding, therefore, are hierarchically related, with counting prerequisite to adding.

Now if we assume that addition is prerequisite to a still more complex task--for example, some forms of multiplication--and that the multiplication task is prerequisite to a still more complex task, we have a hierarchically organized sequence of tasks.

A hierarchy of learning tasks need not be linear. In the hypothesized hierarchy shown in Figure 1, for example, task D is prerequisite to I, J, and K; while K and L are jointly prerequisite to M. The sequences A-C-D and E-F-G-L are shown as independent of one another.

Two tasks can also be said to be hierarchically related when a) one task is more difficult to perform than the other, and b) anyone who can perform the more complex task can reliably be expected to perform the simpler one. For the branching sequence shown in Figure 1, this definition of a hierarchy would require that all subjects who passed a test for objective M also passed the tests for objectives K and L. Anyone who passed L must pass G, anyone who passed G must pass F and E, and so on.

As might be expected, this second definition of a hierarchy has greatly interested testing and evaluation specialists, particularly those concerned with designing diagnostic or placement tests for individualized educational programs. Once hierarchical sequences of this kind have been empirically validated, highly efficient testing procedures can be developed in

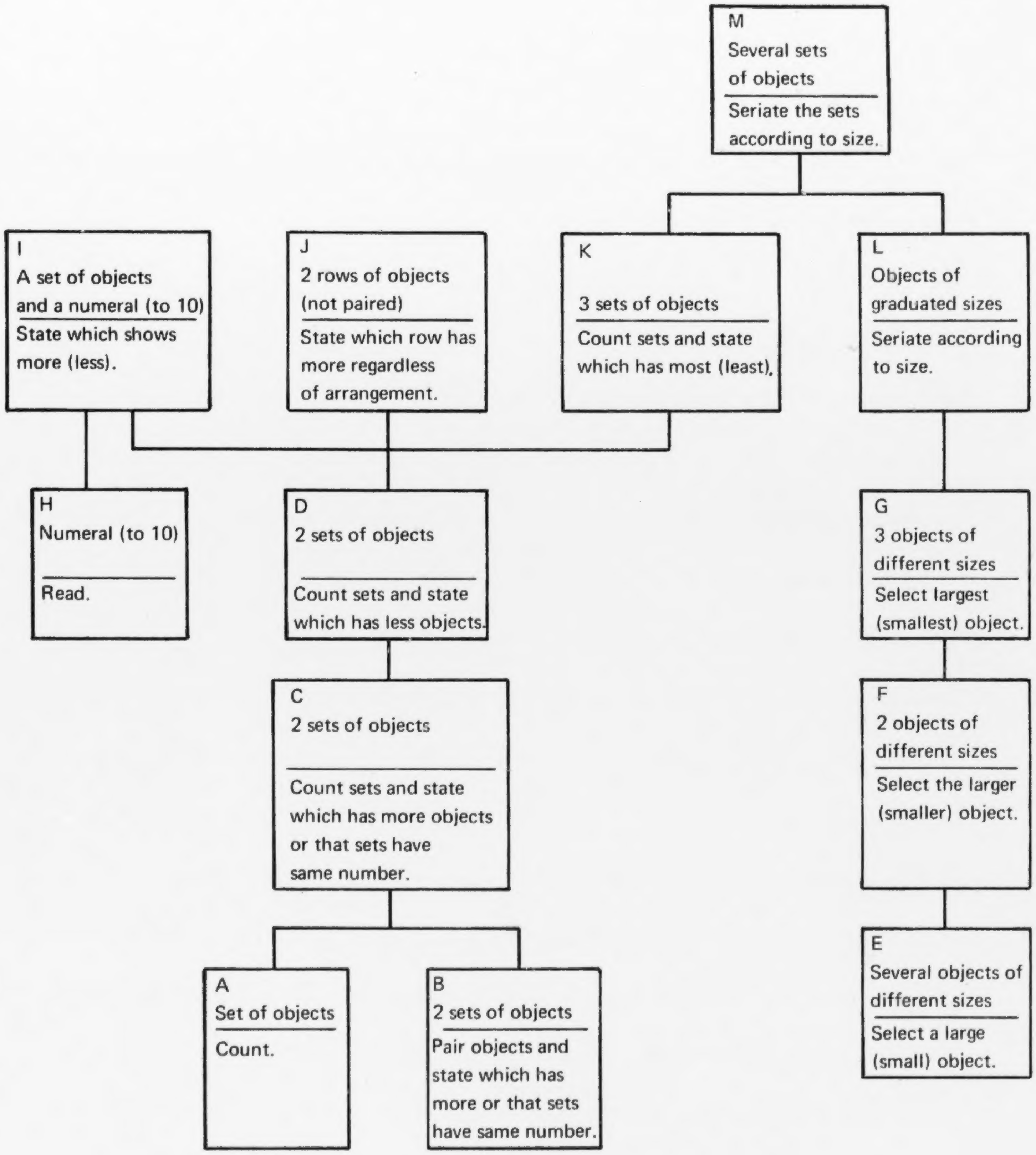

M
Several sets of objects
Seriate the sets according to size.
I
A set of objects and a numeral (to 10)
State which shows more (less).
J
2 rows of objects (not paired)
State which row has more regardless of arrangement.
K
3 sets of objects
Count sets and state which has most (least).
L
Objects of graduated sizes
Seriate according to size.
H
Numeral (to 10)
Read.
D
2 sets of objects
Count sets and state which has less objects.
G
3 objects of different sizes
Select largest (smallest) object.
C
2 sets of objects
Count sets and state which has more objects or that sets have same number.
F
2 objects of different sizes
Select the larger (smaller) object.
A
Set of objects
Count.
B
2 sets of objects
Pair objects and state which has more or that sets have same number.
E
Several objects of different sizes
Select a large (small) object.

Figure 1: A Hierarchy of Number Concept Tasks
(Adapted from Resnick, Wang, & Kaplan, 1970, Figures 5, 6, and 28.)

which students are first tested at key points in the hierarchy, and lower level tests are given only when failure of a higher level test indicates that a given student lacks some prerequisite. Testing programs based on hierarchies have been proposed by a number of authors, one of whom (Ferguson, 1970) would seek to use the on-line decision-making capability of the computer to "branch" students to appropriate tests in the hierarchy on the basis of their immediate performance. Dr. Wang's paper in this symposium will describe a portion of a program of research engaged in developing an integrated set of empirically validated hierarchies and a diagnostic testing program based on these hierarchies.

Developmental psychologists have employed the concept of hierarchy to explain the occurrence of invariant sequences in the acquisition of concepts and logical structures as well as in physical and psychosocial development (see White, 1965). "Stage" theories of development, such as Piaget's, are hierarchical theories in that they propose that an individual can reach a higher stage of development only by passing through a fixed series of lower stages. When such theories stress the orderly sequence of development, rather than specific ages at which particular behaviors are acquired, their hierarchical character becomes especially evident. Such theories essentially predict the order in which certain behaviors will appear. They need not necessarily imply a "maturational" as opposed to learning or organism-environment "interaction" theory of how such changes occur (Spiker, 1966).

Stage theories typically involve relatively large units of analysis, with each unit subsuming a wide variety of specific behaviors. In Piagetian theory, for example, the preoperational stage and the stage of concrete operations can be thought of as constituting two levels of a hierarchy; each level is defined in terms of a range of specific cognitive behaviors, all or most of which must be present for a child to be judged "in" that level or stage. Invariant sequences of cognitive development have also been proposed for much more finely scaled units of behavior, however. For example, Wohlwill (1960) studied the sequence of a set of closely related number behaviors leading to conservation of number; and Smedslund (1964) studied the sequence of a number of specific classification, seriation, and number behaviors, all of which typically appear between six and eight years of age.

Generation of Hypothesized Hierarchies

In his initial article on learning hierarchies, Gagné (1962) proposed that "subordinate learning sets" for a given "terminal" behavior could be generated by asking the single question, "What kind of capability would an individual have to possess to be able to perform this task successfully, were we to give him only instructions?" One or more subordinate tasks are specified in response to this question. The question is then applied to the subordinate tasks themselves, and so on successively down the hierarchy until tasks that can be reasonably assumed in the student population are identified.

Although the logic of identifying prerequisites is clear enough in Gagné's work, he offers few guidelines as to precisely how to identify the "kinds of capabilities" needed in order to learn a given task. It would be difficult, for example, to specify how to train someone in task analysis using Gagné's guidelines alone. A more rigorous technique of analysis focuses on specifying the chain of component behaviors comprising skilled performance and then seeking prerequisites (or "sub-chains") for those components. The method has proved particularly useful to us in sequencing closely related sets of cognitive skills and sometimes in identifying alternative solutions to a given problem which, in turn, yield somewhat different prerequisites.

The task of seriating objects by size provides an interesting example. Seriation is one of the skills generally thought to mark entrance into the stage of concrete operations--generally at seven to eight years of age. Inhelder and Piaget (1964), in their study of seriation, have described a set of partial or "trial and error" solutions which precede "operational" seriation in which the largest, then the next largest, and so on, are systematically selected. Our behavior analysis suggests that a skilled seriator might use either of two methods. An analysis of the first, which Piaget does not describe directly, is shown in Figure 2.

The top box in the chart describes the task in behavioral terms: "Given objects of graduated sizes, the child can seriate according to size."

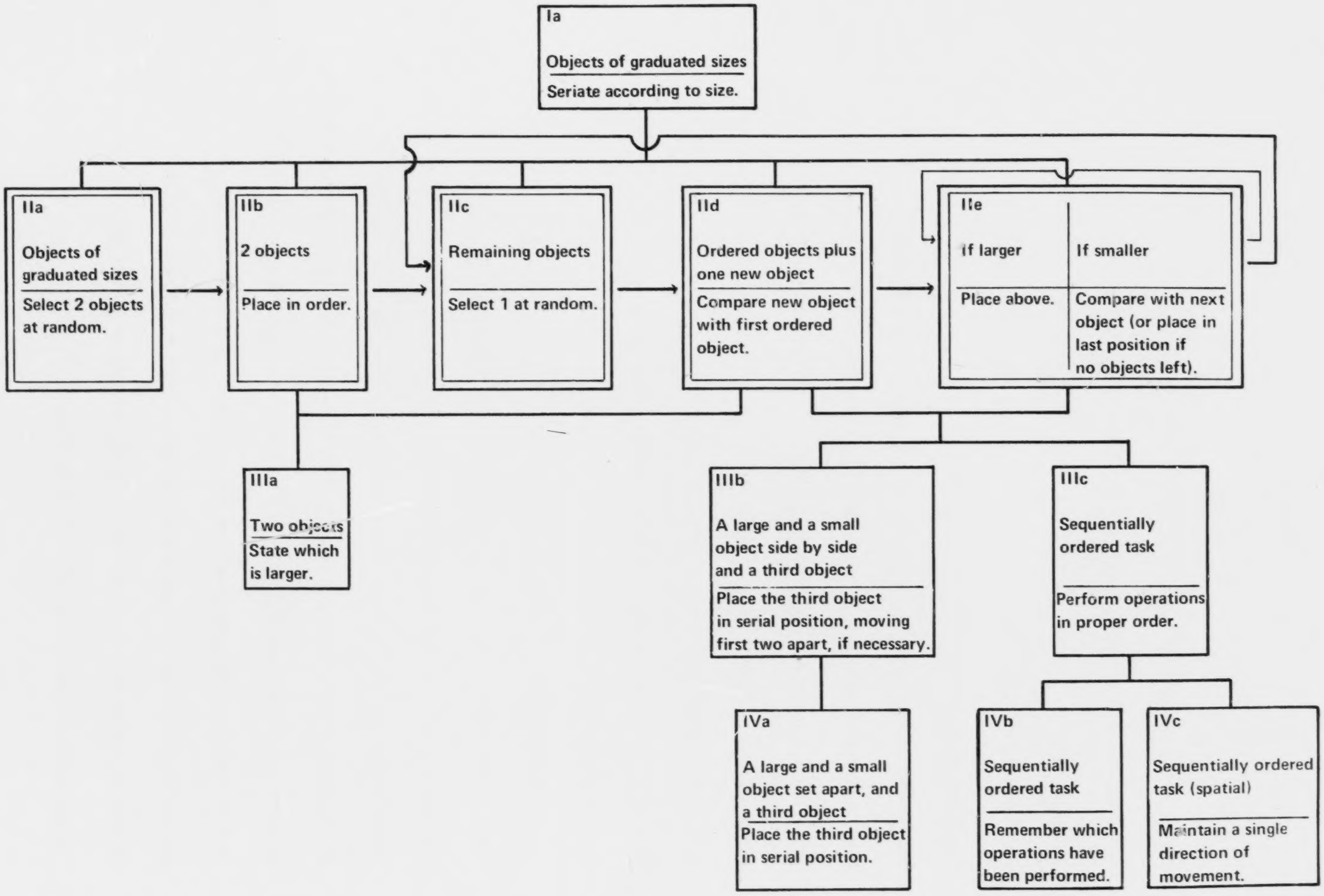

Ia
Objects of graduated sizes
Seriate according to size.
IIa
Objects of graduated sizes
Select 2 objects at random.
IIb
2 objects
Place in order.
IIc
Remaining objects
Select 1 at random.
IId
Ordered objects plus one new object
Compare new object with first ordered object.
IIe
If larger
If smaller
Place above.
Compare with next object (or place in last position if no objects left).
IIIa
Two objects
State which is larger.
IIIb
A large and a small object side by side and a third object
Place the third object in serial position, moving first two apart, if necessary.
IIIc
Sequentially ordered task
Perform operations in proper order.
IVa
A large and a small object set apart, and a third object
Place the third object in serial position.
IVb
Sequentially ordered task
Remember which operations have been performed.
IVc
Sequentially ordered task (spatial)
Maintain a single direction of movement.

Figure 2: Analysis of the Size Seriation Task (Method 1).

The next line shows the chain of component behaviors. The child first selects two objects at random and places them in order (IIa and IIb). He then selects another object at random (IIc) and compares it with the first ordered object (IId). If it is larger, it is placed above the first object; if smaller, a comparison with the next object is made (IIe) until the proper place is found. The chain of selecting and comparing (IIc-IIe) is then repeated for all remaining objects in the set. Hypothesized prerequisites for these components are shown in the lower lines of the chart. They include the ability to state which of two objects is larger (IIIa), to insert a third object in serial position with respect to two already ordered objects (IIIb), and to "keep one's place," remembering which operations have been performed and in which spatial direction one is moving (IIIc, IVb, IVc).

Analysis of a second solution to the seriation task is shown in Figure 3. This is essentially Piaget's "operational" seriation method. The component chain in the second line shows a process in which the child selects the largest object in the set (IIa) and places it in the first position (IIb). He then selects the largest of the remaining objects (IIc) and places it (IId), and then recycles through this chain until all objects have been selected. Some prerequisites are shared with the first solution (see IIIa and IIIc). In addition, this solution requires the ability to recognize and correct errors in placement (IIIb). Thus, in this case behavior analysis procedures specify as a prerequisite one of the "preoperational solutions" described by Piaget.

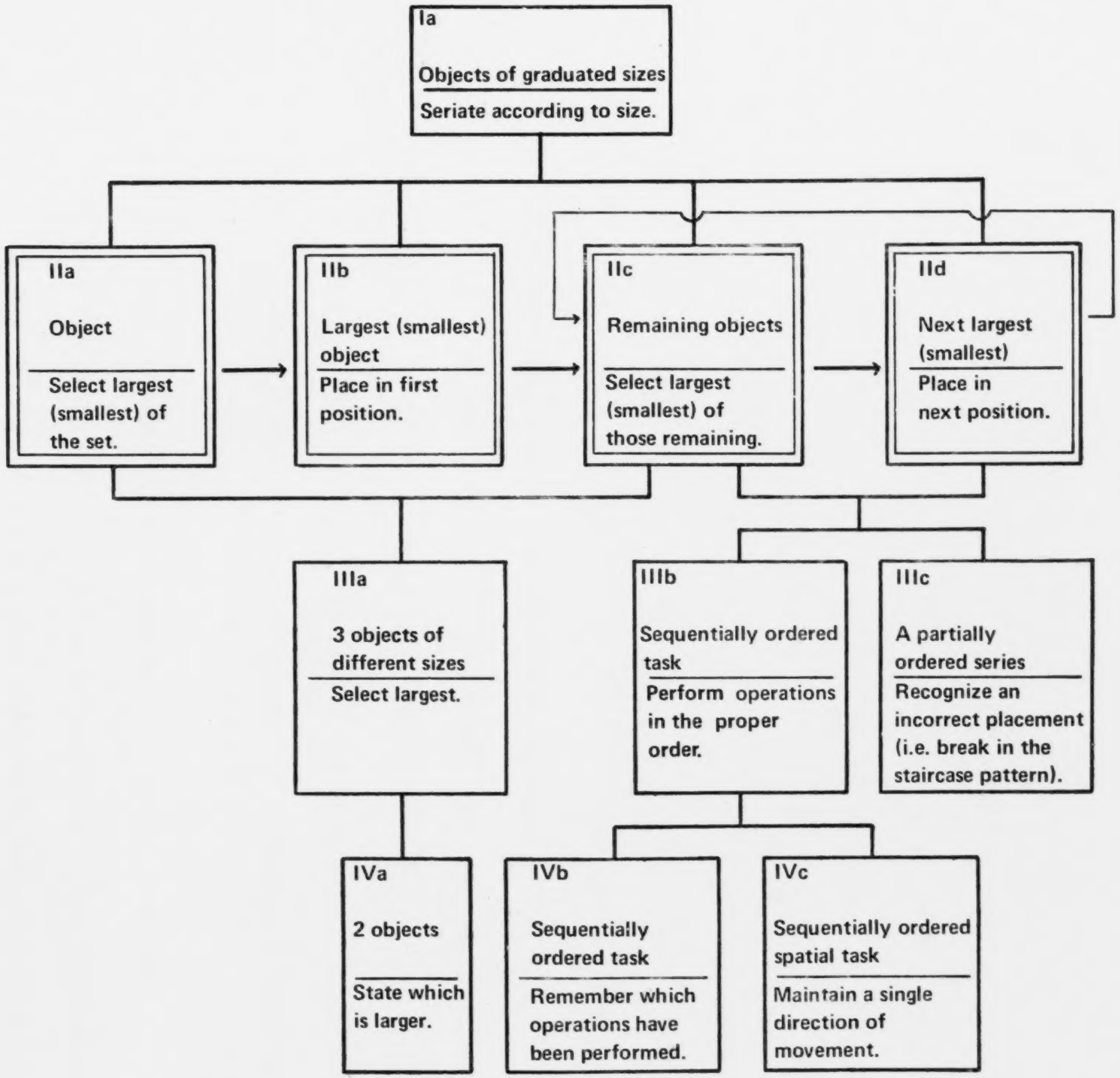

Ia
Objects of graduated sizes
Seriate according to size.
IIa
Object
Select largest (smallest) of the set.
IIb
Largest (smallest) object
Place in first position.
IIc
Remaining objects
Select largest (smallest) of those remaining.
IId
Next largest (smallest)
Place in next position.
IIIa
3 objects of different sizes
Select largest.
IIIb
Sequentially ordered task
Perform operations in the proper order.
IIIc
A partially ordered series
Recognize an incorrect placement (i.e. break in the staircase pattern).
IVa
2 objects
State which is larger.
IVb
Sequentially ordered task
Remember which operations have been performed.
IVc
Sequentially ordered spatial task
Maintain a single direction of movement.

Figure 3: Analysis of the Size Seriation Task (Method 2).

Component analyses of a set of interrelated tasks can provide both a basis for sequencing tasks in an instructional program and suggestions for special teaching strategies for children who have difficulty in acquiring particular tasks. The detailed analyses serve to identify key behaviors which underlie the stated objectives and may need to be taught explicitly to some children. Often, two superficially similar tasks differ with respect to their demands on some basic function such as memory or perceptual organization. These differences between tasks, together with information concerning the shared components which would facilitate transfer, provide the bases for ordering tasks according to complexity and thus for predicting optimal instructional sequences.

It should be noted that this method of analysis into components is limited to behaviors that can be characterized as "chains" or procedures. Hierarchies of verbal knowledge could not be generated in this way. Indeed, it seems unlikely that hierarchical analysis--in the sense of prerequisites and learning sequences--is applicable to the problem of acquiring a body of knowledge.

Like other forms of task analysis, the technique of specifying component behaviors in a chain focuses attention on the characteristics of skilled performance of the task, and the prerequisites for such performance. There is, theoretically at least, no need to consider the characteristics of unskilled performance, except possibly to alert instructional designers to the types of learning "errors" they may need to counteract. Most developmental psychologists, by contrast, are likely to want to attend to the performance of children--i.e., unskilled individuals--at various points in their development and to generate hypotheses concerning developmental sequence from these observations. Thus, most of the scalogram studies of intellectual development, as well as many cross-sectional studies, base their hypotheses on Piaget's naturalistic descriptions of the intellectual behavior of children at different ages. Hierarchies generated in this way will tend to treat errors and misunderstandings typical of earlier stages as necessary prerequisites for reaching higher stages, although task analysis procedures might suggest that the earlier level of performance could be avoided entirely if direct instruction in the components of skilled performance were given.

Validation of Hypothesized Hierarchies

Methods of empirical validation of hierarchies have differed according to the definition of a learning hierarchy accepted by a given investigator and the use he intends to make of his findings. There are three basic validation strategies that have been used, each of which will be represented in the studies to be reported in this symposium.

"Psychometric" validation procedures are those in which a battery of tests sampling the various behaviors in a hypothesized hierarchy is administered and relationships among test scores are examined for dependency relationships--i.e., the extent to which passing one test reliably predicts passing another. Various statistics have been proposed for assessing the degree of dependency between tests. The basic datum for virtually all of these tests can be expressed as a four-fold contingency table, showing pass-fail relationships for pairs of objectives. Table 1 shows a hypothetical table for two objectives. Nineteen Ss passed both objectives; twelve failed both objectives. Eleven Ss passed objective A, but failed objective B. No Ss passed B while failing A. Based on these data, objective B can be said to "depend on" objective A, since everyone who passed B also passed A. Procedures such as scalogram analysis (Guttman, 1944) examine the pass-fail contingencies of an entire set of tests to determine whether the set can be organized hierarchically--i.e., in a nesting sequence with each successive test depending on all previous tests.

Psychometric validation procedures are most directly suited to the purposes of educational evaluation and measurement specialists seeking to develop efficient placement and diagnostic testing procedures. They have also been used by developmental psychologists in studies of the sequence of acquisition of early cognitive skills (e.g., Kofsky, 1963; Wohlwill, 1960). In such studies the sequence of acquisition is inferred from the fact that any child who can perform higher level behaviors can also perform all lower level behaviors in the hierarchical sequence. This represents a more powerful assessment of sequence than the traditional cross-sectional design, in which different groups of children of different ages are compared for performance on a set of tasks.

TABLE 1

Hypothetical Pass-Fail Contingency Table for Two Tasks

		Objective A	
		Pass	Fail
Objective B	Pass	19	0
	Fail	11	12

However, the sequences cannot be as definitive as those observed directly--for example in longitudinal studies, or in training studies where new behaviors are acquired in the course of the experiment itself.

Since they involve no direct instructional intervention, psychometric validations are only suggestive with respect to the transfer properties of hypothesized hierarchies. Specifically, dependency relations between two objectives show that under existing cultural or educational conditions one task is normally learned before the other. Such a relationship does not show either that there is direct transfer from the earlier to the later learned objective or that the dominant order of acquisition is the most efficient.

To determine transfer relationships and optimal teaching orders, studies involving direct instructional intervention are required. The most direct tests of transfer relationships occur when two or more tasks are taught in hypothesized optimal and non-optimal orders. Effects of these orders on the rate of learning individual tasks and of acquiring the whole set of tasks can then be examined. In the present symposium, Dr. Siegel's and Dr. Uprichard's papers describe transfer studies of this kind. Both examine the relationships among closely similar tasks and are able to establish both optimal teaching sequences and clear transfer relationships. Such studies have clear relevance to the design of instructional programs. With respect to developmental theory, such studies, by demonstrating that the acquisition of specific prerequisite behaviors establishes conditions under which relatively complex cognitive skills can be learned, offer a potential challenge to the now dominant view that "general experience" (Kohlberg, 1968) rather than specific learning is necessary for a child to move to a higher stage of cognitive development.

Dr. Wiegand's study poses a similar challenge to the cognitive-developmental point of view, showing that acquisition of "subordinate" tasks leads to acquisition not only of the terminal task itself, but of a logically similar "transfer task." Methodologically, Dr. Wiegand's study uses a combination of psychometric and instructional procedures. Her method is an adaptation of the one used by Gagné and his associates (Gagné, 1962; Gagné, Mayor, Garstens, & Paradise, 1962; Gagné & Paradise, 1961) in their original hierarchy studies. In Gagné's studies, students typically worked through a teaching program designed to teach each objective in a hierarchy in the hypothesized optimal sequence. Upon completion of the program, subjects were tested on mastery of each separate behavior, as in psychometric studies. Dr. Wiegand's study substitutes for the teaching program the opportunity to simply take tests on the behaviors in the hierarchy, moving in an order from simplest to most complex, and thus facilitating "recall" of successive prerequisites.

Hierarchy Theory, Learning, and Individual Differences

It is frequently objected that theories of learning hierarchies necessarily imply that all learning proceeds in small carefully graded steps, and that a single sequence of steps is followed by all individuals. Hierarchy theory, it is implied, cannot account for sudden leaps in learning nor for the variable patterns of learning displayed by different individuals. Those concerns are serious enough to warrant some consideration in this context.

You may recognize in the first objection a rephrasing of the issue of "programmed" versus "discovery" learning. Establishing hierarchies of behaviorally defined learning objectives can be thought of as a means of identifying the small steps called for in theories of programmed instruction. To teach the task at the top of the hierarchy with the fewest errors, one would begin with the simplest task and work up through the successively more complex tasks, taking advantage of positive transfer between tasks. The more finely graded the sequence--i. e., the more completely prerequisites for each successive task had been identified--the more nearly errorless would be the learning. Discovery learning, by contrast, would take place when an individual skipped over prerequisites and moved directly to the top of the hierarchy.

It seems quite evident that there are indeed occasions when individuals actually or apparently learn complex behaviors without explicitly practicing prerequisites for those behaviors. Such occasions may arise, for example, when motivation to learn is high, when only a small number of prerequisites are missing from the individual's repertoire, and when the task is structured in such a way as to suggest the applicability of skills and concepts already in the individual's repertoire. On other occasions, all of the component behaviors for a new task may be present in the repertoire, needing only to be combined and organized into a complex new performance. On such occasions, direct practice on the terminal might produce dramatic learning effects, together with the impression of skipping prerequisites in a hierarchy.

It should be noted that hierarchies of behaviors may be valid even if individuals on occasion do not learn the behaviors in the hypothesized order. Both Siegel and Uprichard, for example, will report some data in which subjects who learn a complex task first, without explicit practice on prerequisite tasks, show evidence of having acquired a simpler, prerequisite task in the process. In these studies, some (and only some) subjects learned in a nonhierarchical sequence, but they nevertheless acquired on their own the prerequisites specified in the hierarchy. Thus the hierarchies appear to be valid, although some subjects did not need the support of an optimally sequenced learning program.

Just as there may be conditions conducive to sequential learning and other conditions that encourage leaping to the top of a presumed hierarchical sequence, so there may be individuals who characteristically learn in small steps and others who can usually take larger jumps, acquiring necessary subordinate skills almost tangentially. Such a difference in learning style, in fact, may well distinguish highly intelligent from less intelligent individuals. I am unaware of any existing research on this aspect of learning style. However, should stable individual differences in the ability to skip over prerequisites in the course of learning be demonstrated, it would be of considerable theoretical and applied interest to then search for the kinds of generalized competencies--linguistic, attentional, imitative, for example--that mediate this ability.

The possibility of alternate routes to a given learning objective can be handled within hierarchy theory by the concept of branching. When a terminal task has two or more prerequisites that are independent of one another, the prerequisites can be learned in any order as long as all have been acquired by the time the terminal is learned. "Disjunctive branches" provide for still more variability in sequence. Disjunctive branches are points in a hierarchy where either one subordinate capability or another would be sufficient for acquiring a new behavior; both are not needed. For example, where there are alternate solutions to a problem, such as the ones shown earlier for seriation, acquiring either set of prerequisites should make it possible to learn the target task. Research to determine the frequency of occurrence of such disjunctive branches is still required.

It should be pointed out, however, that hierarchy theory does indeed imply that for any given set of learning tasks the number of alternate sequences will be limited. While occasional disjunctive branches can be usefully accommodated in a hierarchical structure, if many alternate sequences are required to account for the learning patterns observed, then the hypothesis of hierarchical dependency for the particular tasks involved must be rejected. Validation of a hierarchy, by any of the methods discussed earlier, implies that most individuals learn in the limited number of sequences shown in the hierarchy. Validation, in other words, demonstrates the existence of optimal or dominant sequences of acquisition; and such sequences are what we mean when we use the term "learning hierarchy."

References

Ferguson, R. A model for computer assisted criterion-referenced measurement. Education, 1970, 81 (1), 25-31.

Gagné, R. M. The acquisition of knowledge. Psychological Review, 1962, 69, 355-365.

Gagné, R. M., Mayor, J. R., Garstens, H. L., & Paradise, N. E. Factors in acquiring knowledge of a mathematical task. Psychological Monographs, 1962, 76 (Whole No. 526).

Gagné, R. M. & Paradise, N. E. Abilities and learning sets in knowledge acquisition. Psychological Monographs, 1961, 75 (Whole No. 518).

Guttman, L. A basis for scaling quantitative data. American Sociological Review, 1944, 9, 139-150.

Inhelder, B. & Piaget, J. The early growth of logic in the child. New York: W. W. Norton, 1964.

Kofsky, E. A scalogram study of classificatory development. Child Development, 1963, 34, 191-204.

Kohlberg, L. Early education: A cognitive-developmental view. Child Development, 1968, 39, 1013-1062.

Resnick, L. B., Wang, M. C. & Kaplan, J. Behavior analysis in curriculum design: A hierarchically sequenced introductory mathematics curriculum. Pittsburgh: Learning Research and Development Center, University of Pittsburgh, 1970. (Monograph 2)

Smedslund, J. Concrete reasoning: A study of intellectual development. Monographs of the Society for Research in Child Development, 1964, 29, (2, Whole No. 93).

Spiker, C. C. The concept of development: Relevant and irrelevant issues. Monographs of the Society for Research in Child Development, 1966, 31, (5, Whole No. 107).

White, S. H. Evidence for a hierarchical arrangement of learning processes. Advances in Child Development and Behavior. New York: Academic Press, 1965.

Wohlwill, J. F. A study of the development of number by scalogram analysis. Journal of Genetic Psychology, 1960, 97, 345-377.

PSYCHOMETRIC STUDIES IN THE VALIDATION OF AN EARLY LEARNING CURRICULUM[1]

Margaret C. Wang

University of Pittsburgh

The study reported in this paper is part of an ongoing research program whose cumulative aim is to generate empirically validated hierarchical curriculum sequences in several basic skill areas appropriate to an early learning curriculum. The validated curriculum sequences are used as the basis for a diagnostic testing program through which children can be placed in a curriculum with reasonable assurance that necessary prerequisites have been met.

The particular curriculum sequences to be discussed here are drawn from the 1969-1970 revision of the Primary Education Project's[2] quantification curriculum. The curriculum, designed for pre-school and first-grade children, consists of fourteen units, with each unit made up of a hierarchically arranged sequence of instructional objectives. Figure 1 shows the units of the curriculum in their hypothesized hierarchical sequence. Included in the curriculum units were learning sequences derived from the following sources: a) sequences that

---

[1] The research reported herein was supported by a grant from the Ford Foundation and by the Learning Research and Development Center supported in part as a research and development center by funds from the U.S. Office of Education, Department of Health, Education, and Welfare.

[2] The Primary Education Project (PEP) is a research and development project at the Learning Research and Development Center, University of Pittsburgh. The primary purpose of PEP is to develop an individualized early learning curriculum and a school organization that serves young children in a continuous program beginning at age three and running eventually through the primary grades.

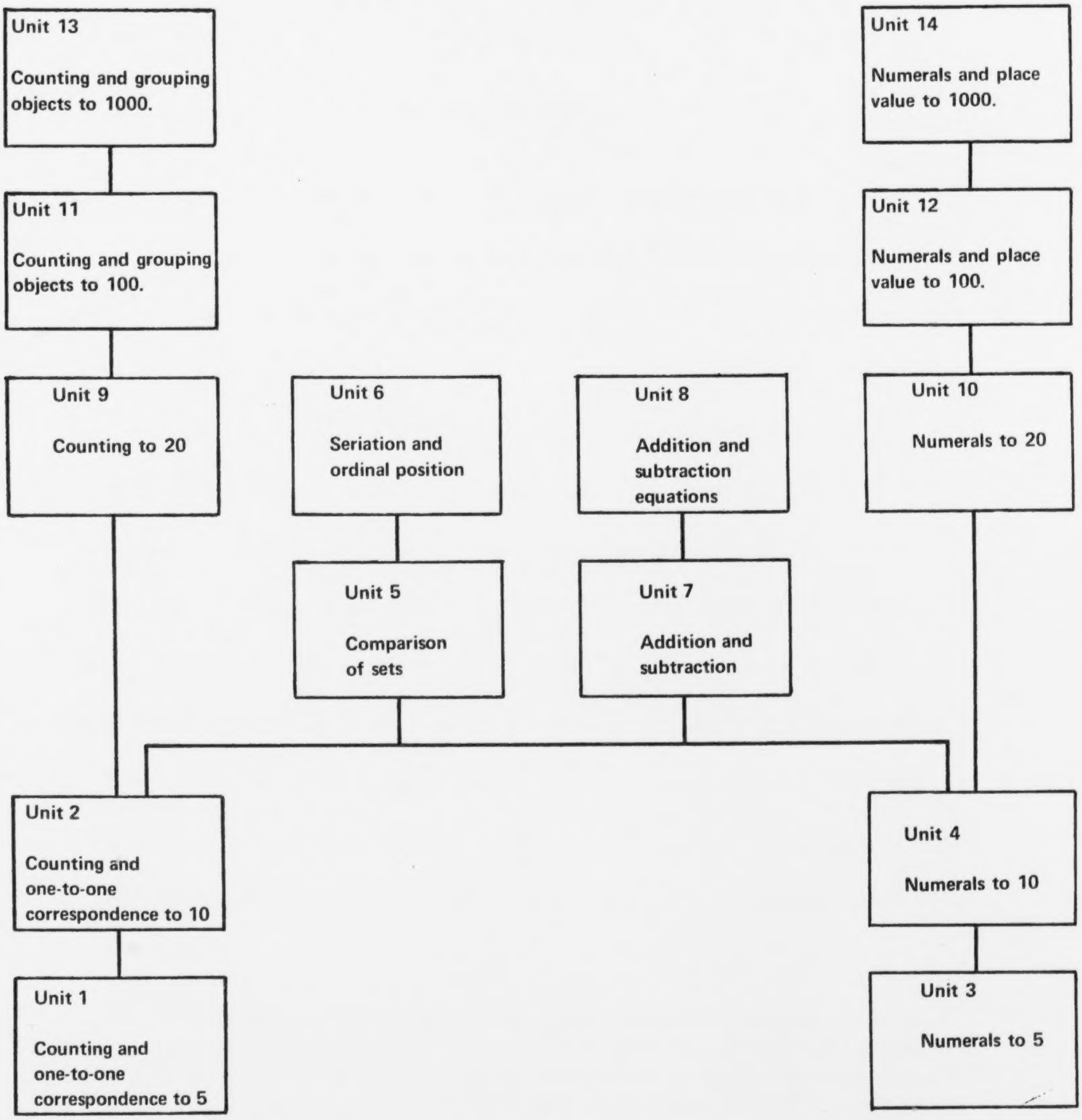

Unit 13
Counting and grouping objects to 1000.
Unit 14
Numerals and place value to 1000.
Unit 11
Counting and grouping objects to 100.
Unit 12
Numerals and place value to 100.
Unit 9
Counting to 20
Unit 6
Seriation and ordinal position
Unit 8
Addition and subtraction equations
Unit 10
Numerals to 20
Unit 5
Comparison of sets
Unit 7
Addition and subtraction
Unit 2
Counting and one-to-one correspondence to 10
Unit 4
Numerals to 10
Unit 1
Counting and one-to-one correspondence to 5
Unit 3
Numerals to 5

Figure 1: Sequence of Units in the PEP Mathematics Curriculum.

had been validated in a previous empirical study (Wang, Resnick, & Boozer, 1971); b) revised sequences suggested by the empirical evidence obtained from the previous study but not yet validated; and c) hypothesized sequences that had not been tested. All sequences were initially derived from a process of detailed behavior analysis, specifying both components and prerequisites of each objective (Resnick, 1967; Resnick, Wang, & Kaplan, 1970).

Two validation questions were asked about these curriculum hierarchies. The first question was addressed to the hierarchical order between the units of instruction; the second question was addressed to the sequential order within each unit. We were, therefore, seeking empirical evidence both for the cumulative dependencies of individual behaviors on one another within a single learning hierarchy, and for the inter-dependencies of a set of learning hierarchies comprising the total curriculum structure.

Method

Subjects

The subjects were pre-kindergarten, kindergarten, and first-grade students in an urban elementary school in Pittsburgh, Pennsylvania. Approximately 90 percent of the children were Black; 49 percent were male and 51 percent were female. A total of 150 children was included in the study. To conserve time and not overburden any child with too many tests, several subsamples were created for different unit validations. The samples for a given validation ranged from 80 to 150. The number of subjects included in each analysis is given in the table of results.

Instrument

Two sets of tests were developed (Wang, 1969): a) a diagnostic test battery which included a test for each individual objective in each unit; and b) a placement test battery which included one test for each unit. Each test consisted of one to five separate items. Children were scored as passing a test only when they passed all items in the test. The tests were designed for individual oral administration to young children. Relevant characteristics of specific tests will be described in conjunction with the interpretation of results.

Procedure

Eight trained research assistants served as testers.[3] The total testing period, which began during the second week of school, lasted about six weeks. During the testing period, the teachers were requested not to use any teaching and learning material in the curriculum in which the children in the class were being tested. However, it should be noted that approximately 37 percent of the kindergarten Ss and 78 percent of the first grade Ss had been enrolled during the preceding year in Primary Education Project classrooms, where they had worked on an earlier version of the PEP quantification curriculum.

The diagnostic tests were administered first. Within each unit, the tests were administered in two different sequences in order to control for learning effects: a) a hypothesized hierarchical sequence beginning with the simplest objectives and moving to the most complex; and b) a random sequence. All the tests in a given unit were administered to each child in the sample regardless of his performance on preceding tests. The battery of placement tests was administered to all children in the sample at the end of the study. These tests were administered in the hypothesized hierarchical order to all subjects.

---

[3] The author wishes to thank John Caruso, Judy Karnas, Francine Landay, Patricia Schuetz, Jane Reynolds, Daniel Rosenthal, Peter Rubinsky, and Maurice Wilson for their assistance in data collection.

Data Analysis

A form of scalogram analysis (Guttman, 1944) was used in analyzing the data. Scalogram analysis provides a method of arranging a set of tests such that passing a test higher in the sequence by any individual reliably predicts passage of all tests lower in the sequence by that individual. Table 1 shows a hypothetical set of perfectly scaled data. Subjects are listed down the side, tests across the top. "0" indicates a failing score, "1" a passing score. Note that once a subject fails a test, he fails all subsequent tests. Conversely, if a subject passes a test, he has passed all earlier tests. It is reasonable to say, therefore, that for any given test in the scale, passing the preceding test is a prerequisite. If any subject were to pass a higher level test and fail a lower level one, the reversal of scores would be counted as an error. In Table 2, two Ss, 1 and 3, are shown with scaling errors. In scalogram analysis the number of errors in a set of data is used to calculate a coefficient of reproducibility, which is a measure of the degree to which the set of data approximates a perfect Guttman scale. A reproducibility of 1.00 indicates a perfect scale.

Multiple Scalogram Analysis (MSA) (Lingoes, 1963) was used in analyzing the present data. This is a method of scalogram analysis which reorders the

TABLE 1

Hypothetical Data for a Perfect Guttman Scale

		Tests				
		1	2	3	4	5
Subjects	1	1	1	1	1	1
	2	1	1	1	1	0
	3	1	1	1	0	0
	4	1	1	0	0	0
	5	1	0	0	0	0
	6	0	0	0	0	0

TABLE 2

Hypothetical Data for a Guttman Scale with Scaling "Errors"

		Tests				
		1	2	3	4	5
Subjects	1	1	1	1	0	1
	2	1	1	1	1	0
	3	1	0	1	0	0
	4	1	1	0	0	0
	5	1	0	0	0	0
	6	0	0	0	0	0

tests in a given battery so as to produce the most optimal scales possible. It will generate one or more linear scales, each one independent of the others--i.e., sharing no tests in common. The hypothesized scale is validated when it matches the empirical scale produced by the MSA program, and is rejected when it differs from the empirical scales. The empirical MSA scale can then be treated as a new hypothesis for future investigations.

For this study the criterion of reproducibility was set at .85. This meant that only those tests that could enter a scale with a reproducibility equal to or greater than .85 were included in a given empirical scale.

Results

Validation of the Between Unit Sequences of the Curriculum

To validate the hypothesized unit sequences the six linear scales implied by the hierarchy shown in Figure 1 were tested using MSA. Separate validation of each linear pathway was required because in any single analysis MSA will produce only independent linear sequences with no objectives in common. Table 3 shows both the hypothesized scales and the empirical scales yielded by MSA. Because very few Ss passed them, units 11, 12, 13, and 14 were not considered in the analysis. In each case the empirical scale is identical to the hypothesized scale. Since the reproducibility coefficient for each empirically validated scale was very high, we can be quite confident of the dependencies shown.

Validation of the Within Unit Hierarchies

In order to empirically validate the hypothesized dependencies of the objectives included in each unit of the curriculum, a separate analysis was performed

TABLE 3

Validation of Between Unit Sequences

N = 150

Hypothesized Scale	Empirical Scale	
	Scale 1	Percent Correct Responses
I		
Unit 13		
Unit 11		
Unit 9	Unit 9	40.3
Unit 2	Unit 2	28.0
Unit 1	Unit 1	18.0
Reproducibility	.969	
II		
Unit 14		
Unit 12		
Unit 10	Unit 10	20.7
Unit 4	Unit 4	12.0
Unit 3	Unit 3	9.3
Reproducibility	.980	
III		
Unit 6	Unit 6	40.3
Unit 5	Unit 5	28.0
Unit 2	Unit 2	4.7
Unit 1	Unit 1	3.3
Reproducibility	.973	
IV		
Unit 8	Unit 8	40.3
Unit 7	Unit 7	28.0
Unit 2	Unit 2	2.0
Unit 1	Unit 1	0.0
Reproducibility	.976	

TABLE 3 continued

Hypothesized Scale	Empirical Scale	
	Scale 1	Percent Correct Responses
V		
Unit 6	Unit 6	20.7
Unit 5	Unit 5	12.0
Unit 4	Unit 4	4.7
Unit 3	Unit 3	3.3
Reproducibility	.983	
VI		
Unit 8	Unit 8	20.7
Unit 7	Unit 7	12.0
Unit 4	Unit 4	2.0
Unit 3	Unit 3	0.0
Reproducibility	.993	

for each unit hierarchy. Because of the limited time available, only a few samples of these analyses are included for discussion here.

Units 1 and 2. Units 1 and 2 cover counting of sets and one-to-one correspondence. Figure 2 shows the hypothesized hierarchical structure of these units. The two units are identical with respect to objectives and hypothesized structure; they differ only in the number of objects to be counted or paired in one-to-one correspondence. Unit 1 covers sets up to five; unit 2 covers sets from six to ten. The hypothesized sequences were tested separately for units 1 and 2, thus providing a form of replication within the study. Table 4 shows predicted and empirical scales for the three linear sequences implied by the hierarchy, shown in Figure 2 (scales A, B, and C).

The hypothesized Scale I included five objectives involving counting objects. Objective A requires only rote counting. B is a basic object counting operation in which the child must synchronize touches with counts. However, since the movable objects can be removed from the set as each is counted, there is no problem of remembering which objects must be counted. Objectives C and D, where objects are fixed, involve this additional memory component. In C the objects are arranged in an orderly pattern; in D they are scattered, posing greater difficulty in remembering which have been counted. In objective F, the child must select among several sets to find one of a stated size. (See Resnick, Wang, & Kaplan, 1970 for a fuller discussion of the rationale for this and other sequence hypotheses.)

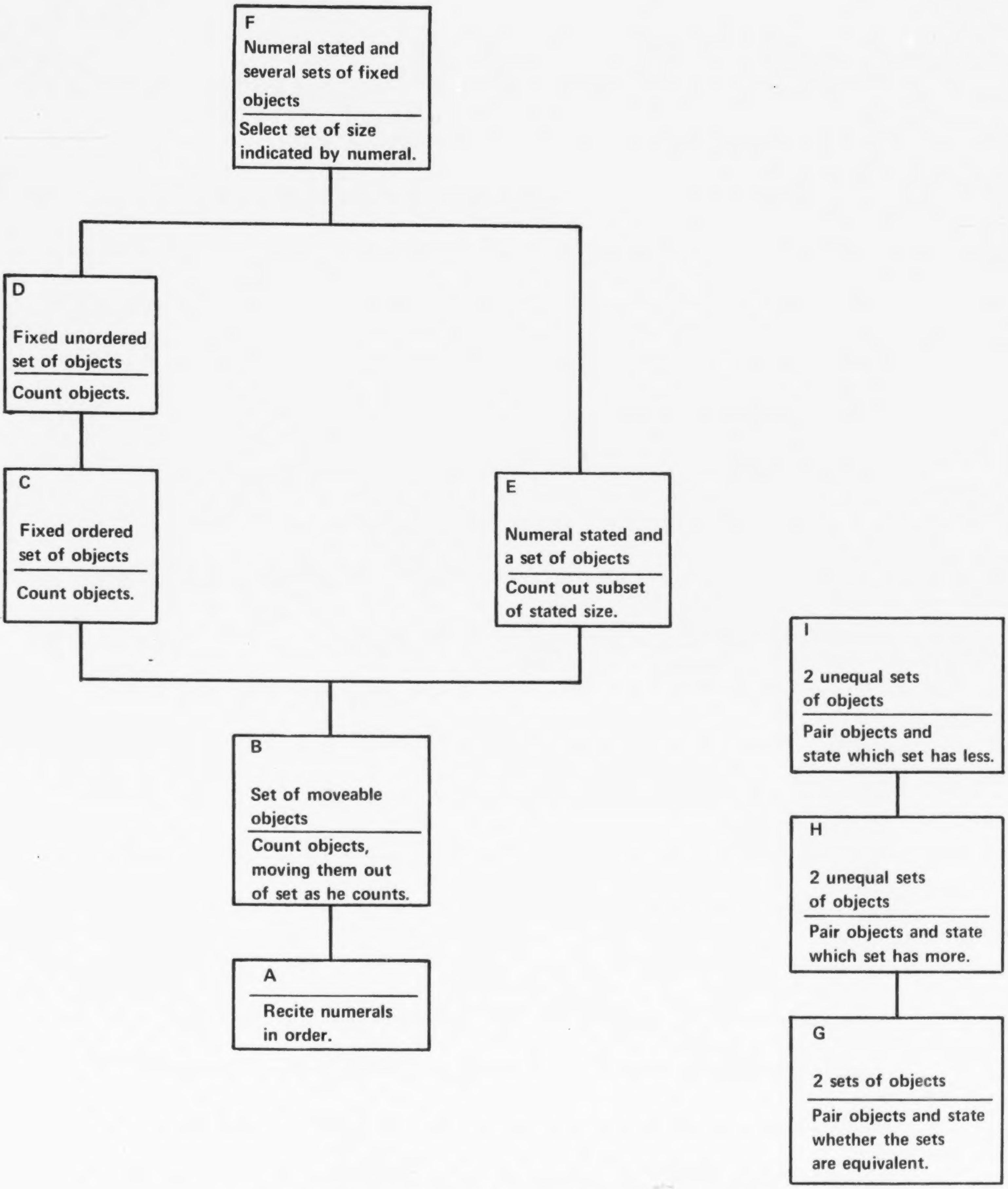

F
Numeral stated and several sets of fixed objects
Select set of size indicated by numeral.
D
Fixed unordered set of objects
Count objects.
C
Fixed ordered set of objects
Count objects.
E
Numeral stated and a set of objects
Count out subset of stated size.
B
Set of moveable objects
Count objects, moving them out of set as he counts.
A
Recite numerals in order.
I
2 unequal sets of objects
Pair objects and state which set has less.
H
2 unequal sets of objects
Pair objects and state which set has more.
G
2 sets of objects
Pair objects and state whether the sets are equivalent.

Figure 2: Hierarchy of Objectives for Units 1 and 2, Counting and One-to-one Correspondence.

TABLE 4

Validation of Within Unit Sequences, Units 1 and 2

N = 82

Hypothesized Scale	Empirical Scales					
	Unit 1		Unit 2			
	Scale 1	% Correct Responses	Scale 1	% Correct Responses	Scale 2	% Correct Responses
Scale I						
Objective F	F	53.7			F	16.3
Objective D	D	53.7			D	13.8
Objective C	B	64.6	B	35.0		
Objective B	C	61.0	C	39.5		
Objective A	A	85.4	A	72.8		
Reproducibility	.934		.959		.969	
Scale II						
Objective F	E	53.7	F	16.3		
Objective E	F	53.7	E	27.5		
Objective B	B	64.6	B	35.0		
Objective A	A	85.4	A	72.8		
Reproducibility	.939		.957			
Scale III						
Objective I	H	23.5	G	35.0		
Objective H	G	42.7	H	18.8		
Objective G	I	72.8	I	35.0		
Reproducibility	.967		.955			

The empirical sequence for Scale I gives only an imperfect match with the hypothesized scale. For unit 2, two independent scales were generated. In both units 1 and 2, objective B, counting movable objects, appears above counting fixed objects. In fact, during testing on objective B, relatively few children took advantage of the possibility of moving objects out of the set. Since E simply spilled the required number of objects onto the table, Ss were, therefore, frequently counting a fixed unordered array (equivalent to objective D). It should be noted that if Ss were explicitly taught to move objects out of the set, they might learn basic counting skills more easily by starting with objective B than with C. Should such a hypothesis be supported experimentally, this would be a case in which scaling and transfer relations of a hierarchy were not identical.

The hypothesized Scale II includes, in addition to rote and movable object counting (objectives A and B), the task of counting out a specified subset from a larger set of objects (objective E) and the task of selecting from among several sets of objects the set that has the stated number of objects (objective F). Objective E requires, in addition to the skill of counting movable objects (objective B), the ability to remember a stated number while counting and to stop at that number. Objective F also requires remembering one number while counting a set; however, the objects are fixed, and presumably, by our prior hypothesis, more difficult to count. Thus, the hypothesized sequence was A-B-E-F. The empirical scales matched the hypothesized scale for the most part, except that in unit 1 the order of objectives E and F is reversed.

Scale III includes three objectives in which the child compares set size by one-to-one correspondence. In the tests for these objectives, the child is asked to pair poker chips of two colors and is then asked whether the rows have the same number of objects (objective G), which row has more (H), or which row has less (I). The predicted sequence (G-H-I) for these objectives was the one proposed by Uprichard in a preliminary report on the experiment he will describe in this symposium. The MSA scales, however, suggest a different sequence, one in which the concept of "more" is learned before equivalence or "same number." It would be of interest to explore the source of this discrepancy. It may lie in the different kinds of tests used in Uprichard's and our studies; or it may represent a case in which scaling and transfer studies yield slightly different hierarchies.

Units 3 and 4. Units 3 and 4 involve recognizing and reading numerals. As for units 1 and 2, the objectives and sequences are identical, with unit 3 covering numerals 1 through 5 and unit 4 covering numerals 6 through 10. The hypothesized hierarchy is shown in Figure 3. The two hypothesized linear sequences were again tested separately for each unit. Table 5 shows the MSA results.

In both units hypothesized and empirical sequences were identical for objectives A, B, and C (this sequence appears in both scales). The same results had been found for these objectives in a study conducted a year earlier with a different sample of children (Wang, Resnick, & Boozer, 1971). Thus, we can have considerable confidence in the sequence which progresses from visual

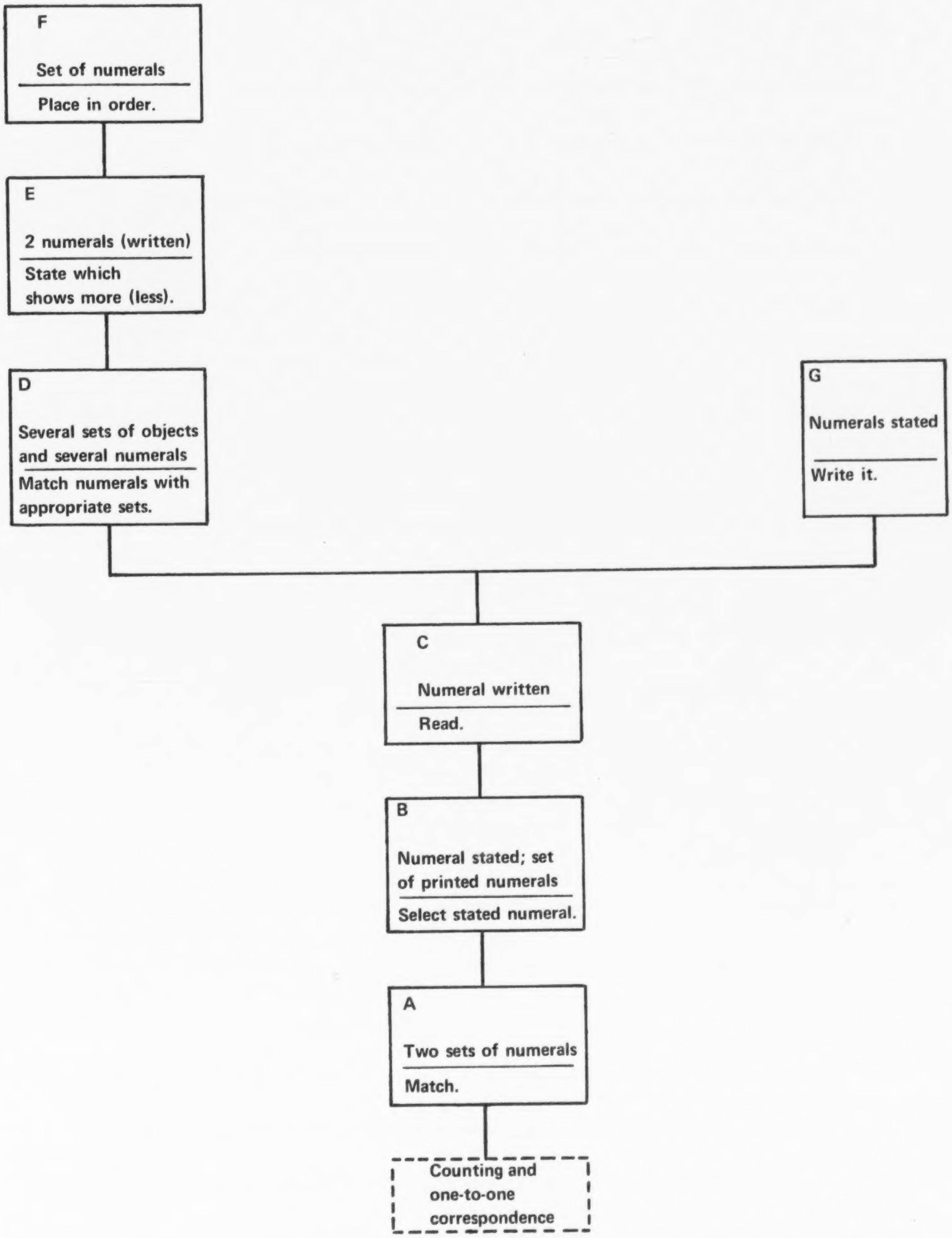

F
Set of numerals
Place in order.
E
2 numerals (written)
State which shows more (less).
D
Several sets of objects and several numerals
Match numerals with appropriate sets.
G
Numerals stated
Write it.
C
Numeral written
Read.
B
Numeral stated; set of printed numerals
Select stated numeral.
A
Two sets of numerals
Match.
Counting and one-to-one correspondence

Figure 3: Hierarchy of Objectives for Units 3 and 4, Numerals.

TABLE 5

Validation of Within Unit Sequences, Units 3 and 4

N = 80

Hypothesized Scale	Empirical Scales			
	Unit 3	% Correct Responses	Unit 4	% Correct Responses
Scale I				
Objective F	E	17.5	D	7.5
Objective E	F	22.5	F	12.5
Objective D	D	26.3	E	12.5
Objective C	C	37.5	C	16.3
Objective B	B	56.3	B	18.8
Objective A	A	85.0	A	95.0
Reproducibility	.965		.977	
Scale II				
Objective G	G	1.3	G	2.5
Objective C	C	37.5	C	16.3
Objective B	B	56.3	B	18.8
Objective A	A	85.0	A	95.0
Reproducibility	.987		.991	

matching (objective A), to identification of numerals when named (objective B), then to naming numerals (objective C). In Scale I, the empirical sequence for objectives D, E, and F is different for the two units, however, and neither agrees with the order in the hypothesized scale. Thus, while MSA confirms the dependence of matching numerals with sets and comparing and sequencing numerals on reading the numerals, it offers no clear information on the relations among the higher level objectives themselves. Empirical results for Scale II match the hypothesized sequence, thus confirming the dependence of skill in writing numerals when named (objective G) on ability to read them. This does not necessarily imply that skill in simply copying numerals would be dependent upon reading them.

Discussion

The data reported here illustrate a way in which psychometric methods can be applied to the problem of validating hypothesized learning sequences. Among the general questions raised by the results is that of the appropriate unit of behavior for hierarchical analysis. It is striking that the hypothesized sequence of curriculum units was directly confirmed in this study, while relations among more finely specified individual objectives were frequently difficult to interpret. In earlier studies (Resnick & Wang, 1969; Wang, Resnick, & Boozer, 1971), the use of single pass or fail scores for groups of closely similar objectives helped considerably in interpreting the data. Findings of this kind suggest that there may be some optimal "unit size" for hierarchy analysis. Determining the characteristics of such optimal units represents a future research problem of considerable importance.

The significance of hierarchy validation studies of the kind we have been reporting lies largely in their relation to a total research program. The establishment of complete curricula with validated hierarchical sequences must inevitably be viewed as a process of successive approximation. With each new study of objectives in a domain, the sequences can be expected to become more refined, and closer matches between predicted and empirical dependencies can be anticipated.

Within this general effort, psychometric validation methods of the kind reported here are most directly relevant to the development of hierarchically based testing batteries. The applicability of psychometric data to instructional design is only indirect, suggesting a likely sequence of acquisition, but not directly testing transfer effects among objectives. Psychometric studies, however, have the advantage of being able to test relationships among a relatively large number of learning objectives within a single study. For this reason, psychometric studies are valuable in organizing a general curriculum area while transfer of training studies are used in studying the relationships among small subsets of objectives.

References

Guttman, L. A basis for scaling quantitative data. American Sociological Review, 1944, 9, 139-150.

Lingoes, J. C. Multiple Scalogram Analysis: A set theoretic model for analyzing dichotomous items. Educational and Psychological Measurement, 1963, 23 (3).

Resnick, L. B. Design of an early learning curriculum. Pittsburgh: Learning Research and Development Center, University of Pittsburgh, 1967. (Working Paper 16)

Resnick, L. B., & Wang, M. C. Approaches to the validation of learning hierarchies. In Proceedings of the Eighteenth Annual Western Regional Conference on Testing Problems. Educational Testing Service, 1969.

Resnick, L. B., Wang, M. C., & Kaplan, J. Behavior analysis in curriculum design: A hierarchically sequenced introductory mathematics curriculum. Pittsburgh: Learning Research and Development Center, University of Pittsburgh, 1970. (Monograph 2)

Wang, M. C. The PEP testing program. Pittsburgh: Learning Research and Development Center, University of Pittsburgh, 1969.

Wang, M. C., Resnick, L. B., & Boozer, R. The sequence of development of some early mathematics behaviors. Child Development, 1971, 42 (6), 1767-1778.

THE EFFECT OF SEQUENCE IN THE ACQUISITION OF THREE SET RELATIONS: AN EXPERIMENT WITH PRESCHOOLERS

A. Edward Uprichard

University of South Florida

By its very nature, mathematics content has an inherent logical structure. An analysis of a mathematics concept or principle would consist of listing prerequisite mathematical learning deemed necessary for the logical development of the given concept. Inherent in the analysis would be hierarchical relationships among various concepts, based on the structure of the discipline. Hence, it is frequently argued, one should follow the structure of the discipline in determining instructional sequences in mathematics.

As a mathematics educator, I am interested in ascertaining whether the structure of mathematics does in fact yield the most efficient instructional sequences for learning. The logical structure of mathematics concepts and the cognitive processes a child goes through in acquiring the concepts may differ. Piaget's work (1968) as well as others' suggests that children's thought processes differ basically from those of adults. Therefore, it is possible that when working with young children, instructional sequences based on the structure of mathematics alone may result in impeding rather than enhancing the cognitive development of children.

The present study was conducted to determine the most efficient instructional sequence through which preschoolers acquire three set relations deemed necessary for the logical development of number: "equivalence," "greater than," and "less than." The efficiency of a particular sequence was evaluated in terms of: a) time to learn the three set relations to criterion, and b) performance on a posttest consisting of both criterion items and transfer items measuring the three relations.

On the basis of mathematical theory it was predicted that the instructional sequences beginning with "equivalence" would be the most efficient. The learning of "equivalence" is thought to be prerequisite to acquiring "greater than" and "less than" since the latter two relations may be mathematically defined in terms of "equivalence." Hence, the following hypotheses were stated:

a) Groups taught "equivalence" first will learn the three set relations fastest. Order in which "greater than" and "less than" are taught will not affect time to reach criterion.

b) On a posttest following instruction, groups taught "equivalence" first will have more correct responses on all three set relations. There will be no differences between the groups taught "greater than" first and those taught "less than" first.

c) On the posttest there will be more correct responses on items testing "equivalence" than on items testing "greater than" or "less than." There will be no difference between the latter two relations.

Method

Independent Variable

Instructional sequence was the single independent variable in this study. There are six possible sequences that can be generated for teaching the set relations "equivalence (E)," "greater than (G)," and "less than (L)": EGL, ELG, GEL, GLE, LEG, and LGE. Each sequence was used as an experimental treatment.

In addition to the sequence treatments, there were two control groups. One control group, C_T, received no instruction but was administered a criterion test weekly (at the same time as the experimental groups) and a posttest at the end of the instructional period. This group's scores were used to establish reliability coefficients between criterion tests. The other control group, $C_{\bar{T}}$, also received no instruction but was administered only the posttest. By comparing posttest results of the two control groups it was possible to determine whether any learning resulted from repeated exposure to the criterion tests.

Dependent Variables

The dependent variable of major interest in this study was the child's ability to identify a set a) "equivalent to," b) "one greater than," and c) "one less than" given sets varying in numerosity and pattern. This ability was measured by a series of tests constructed by the investigator.

The tests consisted of four equivalent criterion tests, Forms A, B, C, D, and a Transfer Test. The criterion tests were administered at one-week intervals. The results of these tests were used to obtain learning curves and to determine the effects of instruction. The last criterion test given (Form D) and the Transfer Test comprised the posttest.

Each criterion test consisted of twelve matching-from-sample tasks on the concrete level: four measuring "equivalence," four "one greater than," and four "one less than." Each task consisted of one sample board and three choice boards. Both the sample board and the choice boards had sets of holes recessed in them. The set of holes in the sample board contained a set of blocks. There were four sets of boards in each criterion test. Each set was used three times, once for each set relation. A child was judged to have reached criterion on a particular set relation if he passed three of the four items for that relation. Figure 1 shows the four sets of boards used in criterion test Form D.

The specific directions given to the child by the examiner when testing the various relations were:

a) Equivalence. E pointed to the sample board and said, "This is a set of three (or four, five, six, as appropriate). Move the blocks to the set that has the same number of holes as this set (referring to the sample board)."

b) Greater Than. E pointed to the sample board and said, "This is a set of three (four, five, six). Move the blocks to the set that has one more hole than three."

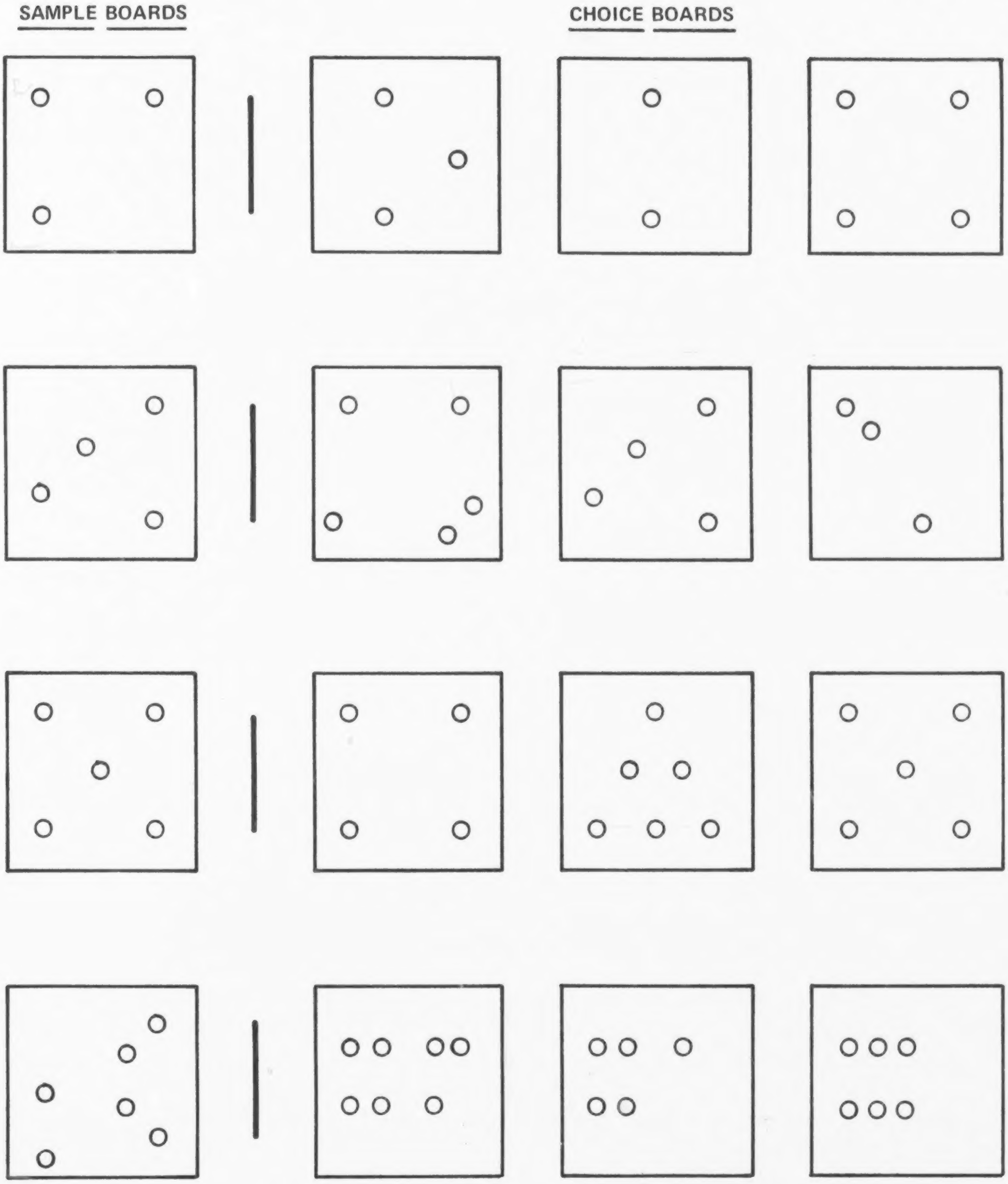

SAMPLE BOARDS
CHOICE BOARDS

Figure 1: Criterion Test, Form D.

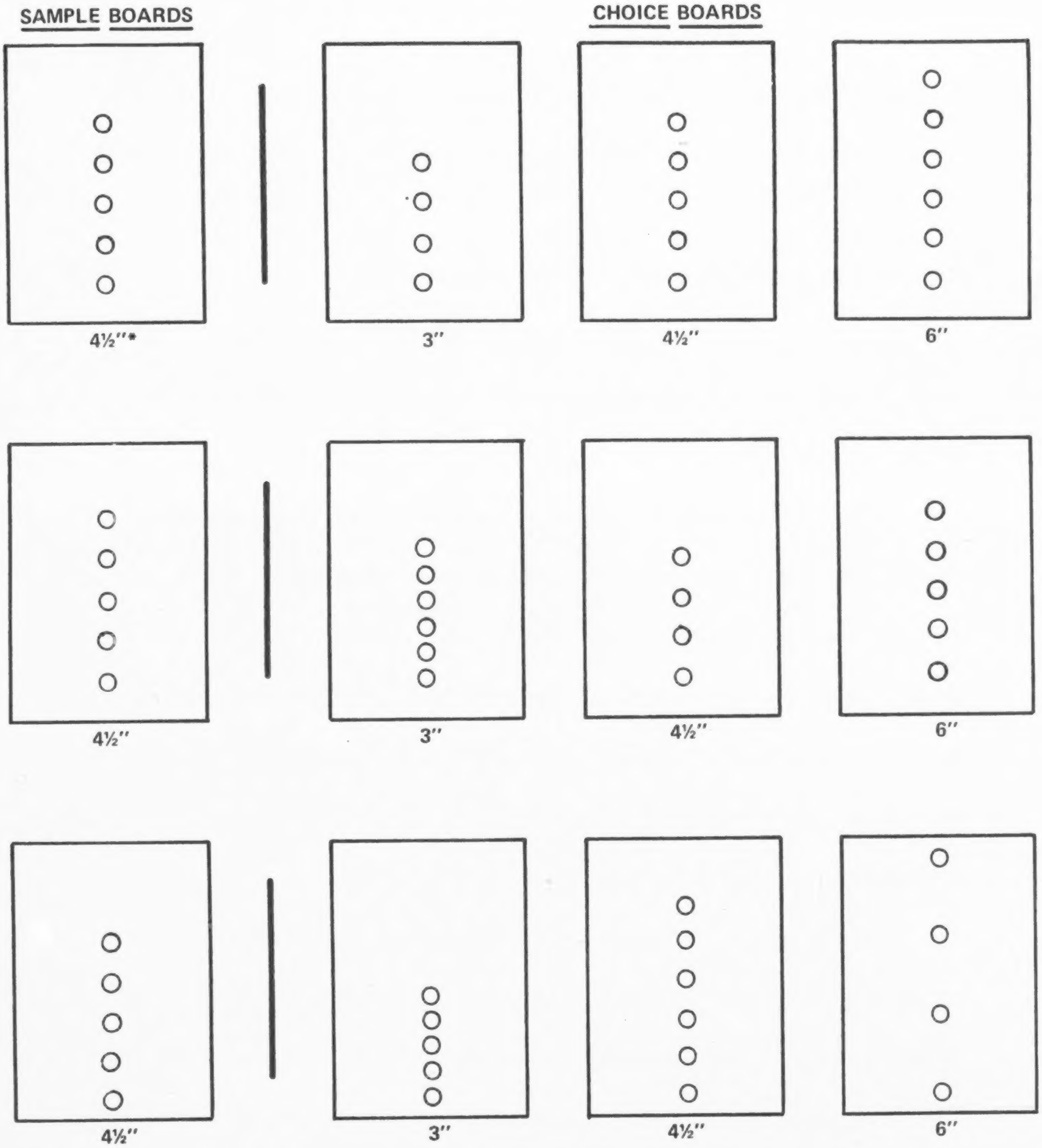
c) Less Than. E pointed to the sample board and said, "This is a set of three. Move the blocks to the set that has one less hole than three."

The Transfer Test administered to the children as part of the posttest also consisted of concrete matching-from-sample tasks: three measuring "equivalence," three "one greater than," and three "one less than." For these tasks only a linear pattern was used, and the only number to be represented on a sample board was five. The three sets of boards used are shown in Figure 2. The testing procedures and directions used with these tasks were identical to those used on the criterion tests. It is interesting to note that the child who used perception alone (length of line) as a strategy for solving these tasks would at most receive credit for three correct responses. This test was classified as a transfer test for two reasons: a) perceptual clues were purposely built into the tasks to distract the children, thus making the test more difficult than the criterion tests; and b) the linear pattern used in this test had not been seen by the children prior to the posttest.

A second dependent variable of interest in this study was time spent in learning. More specifically, the investigator was interested in the amount of time it would take for an experimental group to attain a criterion level on the three set relations. The time variable was measured by the number of weeks of instruction required to meet criterion on each relation.

Subjects and Procedure

A random sample of thirty-two middle-class preschool children ranging in age from four years, one month to five years, one month was


SAMPLE BOARDS
CHOICE BOARDS
4½"*
3"
4½"
6"
4½"
3"
4½"
6"
4½"
3"
4½"
6"

*Linear spread of holes is indicated in inches under each board.

Figure 2: Transfer Test.

randomly assigned to eight treatment groups--the six experimental groups described above and the two controls. There were four subjects in each group. Each experimental group received instruction on the set relations "equivalence," "greater than," and "less than" in one of the six possible sequences. The subjects had approximately twenty-five minutes of instruction each Monday, Tuesday, and Thursday of each week. On Friday of each week, a criterion test was administered to all experimental groups and to the C_T control group.

The four children in each experimental treatment were taught as a group. Two basic activities were used in each instruction session: Identification and Construction. Each activity, presented in a "game-like" approach, could be adapted to teach "equivalence," "greater than," or "less than." No activities used precisely the format or procedure of the criterion test items. Only one set relation was taught to a group in a given week. The criterion and posttests were administered individually.

There was no fixed amount of time allotted for learning a particular set relation in any instructional sequence. The relation taught to an experimental group was determined by the weekly criterion test. If on the criterion test three of the four children in a group reached criterion for the set relation being taught that week, the group was judged to have reached criterion and went on in the following week to receive instruction on the next relation in their particular instructional sequence. Instruction ended for all groups when one group reached criterion on all three set relations. The last criterion test (Form D), administered on Friday, and the Transfer Test, administered the following Monday, comprised the posttest. The C_T control group received both parts of the posttest on Monday.

There were two teachers, two instruction rooms, three periods of instruction (1:00-1:25, 1:30-1:55, 2:00-2:25), and three examiners in this investigation. An attempt was made to randomize the effect upon learning of time and personality of teacher or examiner through systematic rotation. For example, no experimental group received instruction from the same teacher on consecutive instruction days, nor did any group have instruction during the same time period from week to week.

Results

Table 1 shows the number of children in each treatment group who reached criterion on each weekly test. After one week of instruction EGL and ELG were the only groups to reach criterion on the set relation they were learning. At the end of two weeks of instruction both of these groups continued to meet criterion on "equivalence." However, no experimental group attained criterion on the relation on which they received instruction during the second week.

Four groups, EGL, ELG, GEL, and LEG, had reached criterion on "equivalence" after three weeks of instruction on this relation. The GEL group also reached criterion on "greater than" at this time, thus making it the only group to meet criterion on two relations.

TABLE 1

Number of Ss at Criterion on Successive Weekly Tests

Group	Week 1 (Form A)				Week 2 (Form B)				Week 3 (Form C)				Week 4 (Form D)			
	Concept Taught	E	G	L	Concept Taught	E	G	L	Concept Taught	E	G	L	Concept Taught	E	G	L
EGL	E	3*	1	2	G	3*	1	1	G	3*	1	2	G	4*	3*	3*
ELG	E	3*	1	0	L	3*	2	2	L	3*	1	1	L	3*	2	3*
GEL	G	0	1	1	G	2	1	0	G	3*	3*	0	L	4*	3*	0
GLE	G	0	1	0	G	1	2	0	G	2	1	0	G	3*	4*	0
LEG	L	1	0	0	L	2	2	0	L	4*	0	1	L	3*	1	1
LGE	L	0	1	1	L	1	4*	1	L	1	2	0	L	3*	2	1
C_T		2	0	0		0	1	1		1	1	0		0	1	0

*Group at Criterion

On criterion test Form D, administered after four weeks of instruction, the EGL group, which had received instruction on only two relations, "equivalence" and "greater than," reached criterion on all three set relations. The ELG group reached criterion on "equivalence" and "less than," the relations on which they received instruction, and missed meeting criterion on "greater than" by only one correct response. The GEL and GLE groups each reached criterion on two relations, "equivalence" and "greater than." The LEG and LGE groups reached criterion on only one relation, "equivalence." All experimental groups regardless of instructional sequence, attained criterion on "equivalence," thus clearly suggesting the priority of this relation in a learning sequence.

The results support the first part of Hypothesis 1, and the prediction that learning "equivalence" is prerequisite to acquiring "greater than" and "less than." However, they also suggest that "greater than" and "less than" are not simple complements of one another. "Less than" appears considerably more difficult to learn when it is taught first in the sequence.

The existence of a hierarchical relationship among the set relations "equivalence," "greater than," and "less than" is also suggested by the pattern of pass-fail scores for the three relations on the criterion tests. Of 112 individual administrations of criterion tests over a four-week period only eleven instances occurred where children attained criterion on "greater than" or "less than"

56

without doing so on "equivalence." Two children accounted for five of these eleven instances.

Table 2 shows the mean number of correct responses for each relation on the posttest. An analysis of variance test (Table 3) on the combined posttest scores showed significant effects for both the Sequence and Relations factors. Only three of forty-two post-hoc pairwise comparisons tested on the Sequence factor were statistically significant (Scheffe, $\alpha = .10$). These were a) EGL vs. $C_{\bar{T}}$ (F=16.27), b) EGL vs. C_T (F=13.87), and c) the average of EGL and ELG vs. the average of C_T and $C_{\bar{T}}$ (F=25.43). All differences between EGL and the other experimental groups, although not statistically significant, were in the predicted direction. Hypothesis 2 is thus weakly supported with respect to the effect of "equivalence."

All post-hoc pairwise comparisons on the Relations factors were significant (Scheffe, $\alpha = .10$). There were significantly more correct responses on "equivalence" than on "greater than," and more in turn on "greater than" than "less than." These results support the first part of Hypothesis 3, concerning "equivalence," but they contradict the hypothesis that "greater than" and "less than" are of equal difficulty.

Separate analyses of variance were also performed on the two parts of the posttest, criterion test Form D and the Transfer Test. The results exactly paralleled those for the combined scores.

TABLE 2

Mean Number Correct Responses on Each Relation on Posttest

	Equivalence			Greater Than			Less Than			Total
Group	Criterion Test D	Transfer Test	Combined Tests	Criterion Test D	Transfer Test	Combined Tests	Criterion Test D	Transfer Test	Combined Tests	Combined Tests
EGL	3.75	3.00	6.75	3.00	1.25	4.25	2.75	1.25	4.00	5.00
ELG	3.50	2.75	6.25	2.75	0.25	3.00	2.75	1.50	4.25	4.50
GEL	3.25	1.50	4.75	3.25	2.00	5.25	0.75	0.75	1.50	3.83
GLE	2.50	1.50	4.00	3.50	2.00	5.50	0.75	0.00	0.75	3.42
LEG	3.25	1.00	4.25	1.00	1.00	2.00	1.25	1.50	2.75	3.00
LGE	2.75	1.00	3.75	2.75	1.00	3.75	1.50	0.75	2.25	3.25
C_T	1.00	1.50	2.50	2.00	0.75	2.75	0.25	0.50	0.75	2.00
$C_{\overline{T}}$	1.50	1.00	2.50	0.75	1.00	1.75	0.75	0.25	1.00	1.75
Total			4.34			3.53			2.16	

TABLE 3

ANOVA (Two Factor-Repeated Measures) on Mean Number of Correct Responses on Combined Posttest

Source of Variation	SS	df	MS		F
Between Subjects	199.3	31			
Sequence	106.0	7	15		3.85*
Subjects Within Groups	93.3	24	3.9	est. w_A^2=	.38
Within Subjects	244.7	64			
Relations	78.4	2	39.2		19.6*
Sequence X Relation	71.1	14	5.1		2.55*
Relation X Subjects Within Groups	95.2	48	2.0		

*$p < .01$

Discussion

The results of this study suggest that the most efficient sequence for teaching preschoolers the set relations in question is "equivalence," followed by "greater than," then "less than." With respect to the primacy of "equivalence" in the sequence, the empirical sequence and the predicted sequence derived from mathematical theory are in agreement. However, from a purely mathematical point of view, either "greater than" or "less than" might be taught next: there should be no difference in learning difficulty for these two relations since, by definition, if a given set A is greater than a second set B then set B is less than A. The data, in this case, do not support the mathematical prediction, indicating a clear advantage of teaching "greater than" before "less than."

The LEG and LGE groups each received four weeks of instruction on "less than" but failed to reach criterion on this relation. The ELG group learned "less than" after three weeks of instruction on it, after having learned "equivalence" first. The EGL group, after receiving instruction on only "equivalence" and "greater than," reached criterion on all three set relations in the fourth week. These data support the existence of a learning hierarchy of set relations, and further suggest that even a difficult relation such as "less than" can be easily learned when "prerequisite" relations have been learned in an optimal sequence.

It is of interest to note further that the Transfer Test, in particular, tested a form of number conservation since it required a judgment of numerical equivalence (and an actual transfer of the set of blocks) where lengths of the columns were different. Ss in the EGL and ELG group scored correctly on 23 of 24 responses (eight Ss; three responses per S) on the Transfer Test. The two control groups scored only 10 correct of 24 on this test. These results suggest that the training methods used in this study may constitute an effective means of training preschool children to conserve the one-to-one correspondence between equivalent sets.

Reference

Piaget, J. On development of memory and identity. *Heinz Werner Lecture Series*, Vol. 2, 1967, Worcester: Clark University Press.

HIERARCHICAL RELATIONSHIPS AND TRANSFER IN LEARNING MULTIPLICATIVE CLASSIFICATION TASKS[1, 2]

Alexander W. Siegel and Esther Kresh

University of Pittsburgh

Pittsburgh Public Schools

The presence of multiplicative classification skills has generally been considered one index of the child's having reached the stage of concrete logical operations (Flavell, 1963). One of the most interesting manifestations of multiplicative classification skill is the child's ability to deal with two aspects of a situation at a time. A reasonable approach to studying this ability is to examine it in the context of a logically complex classification task, the matrix, which involves the simultaneous ordering of two dimensions. A child who completes or who can construct a double classification matrix is showing some evidence of multiplicative classification ability (Inhelder & Piaget, 1964).

The purpose of this study was to test several hypotheses concerning the nature of hierarchical transfer relationships in two different matrix classification tasks. These tasks require double classification skills; i.e., the child is required to sort on, or deal with, two stimulus dimensions simultaneously. Although different types of matrix tasks have been studied (e.g., Bruner & Kenney, 1966; Lovell, Mitchell, & Everett, 1962; Overton & Brodzinsky, 1969;

---

[1] This study was supported by a grant from the United States Office of Education to the Learning Research and Development Center, University of Pittsburgh. A fuller account of the experiment appears in Resnick, Siegel, and Kresh (1971).

[2] The authors wish to thank Tam Spitzer and John Caruso for their help in running subjects and analyzing data.

Shantz, 1967; Siegel & Kresh, 1971; Smedslund, 1964), most research looks at matrix behavior from a developmental point of view, and seeks to investigate the chronological development of the ability to perform various kinds of matrix tasks. Although variations in the nature of the task would be expected to materially affect performance, few studies have attempted to systematically analyze the behaviors required by the task (Smedslund's papers, 1967a and 1967b, are exceptions). Furthermore, there have apparently been no studies of transfer effects among different matrix tasks.

Gagné's research has shown that when instruction in complex intellectual tasks proceeds upward through a hierarchy of increasingly complex tasks, each one prerequisite to the next, nearly uniform positive transfer from one task to the next occurs (Gagné, 1962; Gagné, Mayor, Garstens, & Paradise, 1962; Gagné & Paradise, 1961). Furthermore, if the subject has learned the prerequisites in order of increasing complexity or difficulty, the terminal task itself can often be "learned" without explicit instruction. None of Gagné's studies of learning hierarchies, however, has directly tested the asymmetry of the transfer effect. That is, the studies were not designed so that the effects of learning the tasks in the hypothesized optimal order could be compared with learning them in non-optimal orders. The present study was specifically designed to make such a comparison for two matrix classification tasks.

Our hypotheses concerning optimal learning order (hierarchical relationships) for two different matrix classification tasks were derived from a systematic behavior analysis of the kind described and illustrated by Resnick in the opening paper in this symposium. Figure 1 shows the stimulus layout for a matrix task in which "attribute cells" explicitly define the objects for each matrix cell. The child's task is to place objects in the appropriate cells according to color (shading) and shape. The behavior analysis of this task, presented in Figure 2, indicated that three steps must be followed in order to place an object in the proper cell. The correct row must be found (Box IIa in Figure 2), the correct column must be found (Box IIb), and then the intersect of the row and column must be found (Box IIc). The process of identifying the appropriate row or column is actually a form of matching-to-sample task in which the object to be placed is the "sample" stimulus and the attribute cells constitute the "choice" stimuli. Box IIIa, therefore, describes matching-to-sample behavior as a prerequisite to both IIa and IIb. Only a relatively simple form of matching-to-sample is required, as the choice stimuli vary in only a single dimension (e.g., color or shape); there is no intruding irrelevant dimension which S must learn to ignore. This restriction is indicated in Box IIIa. A still simpler form of matching-to-sample, in which an identical match is possible, is shown as a lower-level prerequisite (Box IVa).

Once the proper row and column have been identified, finding the intersect is a fairly mechanical matter. However, it does involve certain spatial organization behaviors which permit one to "keep one's place" in a relatively complex visual field. A hypothesized sequence of such spatial organization skills, cumulatively prerequisite to locating the intersect of a row and column, is shown in Boxes VIa, Va, IVb, and IIIb. No linguistic encoding appears necessary to the solution of this task.

85

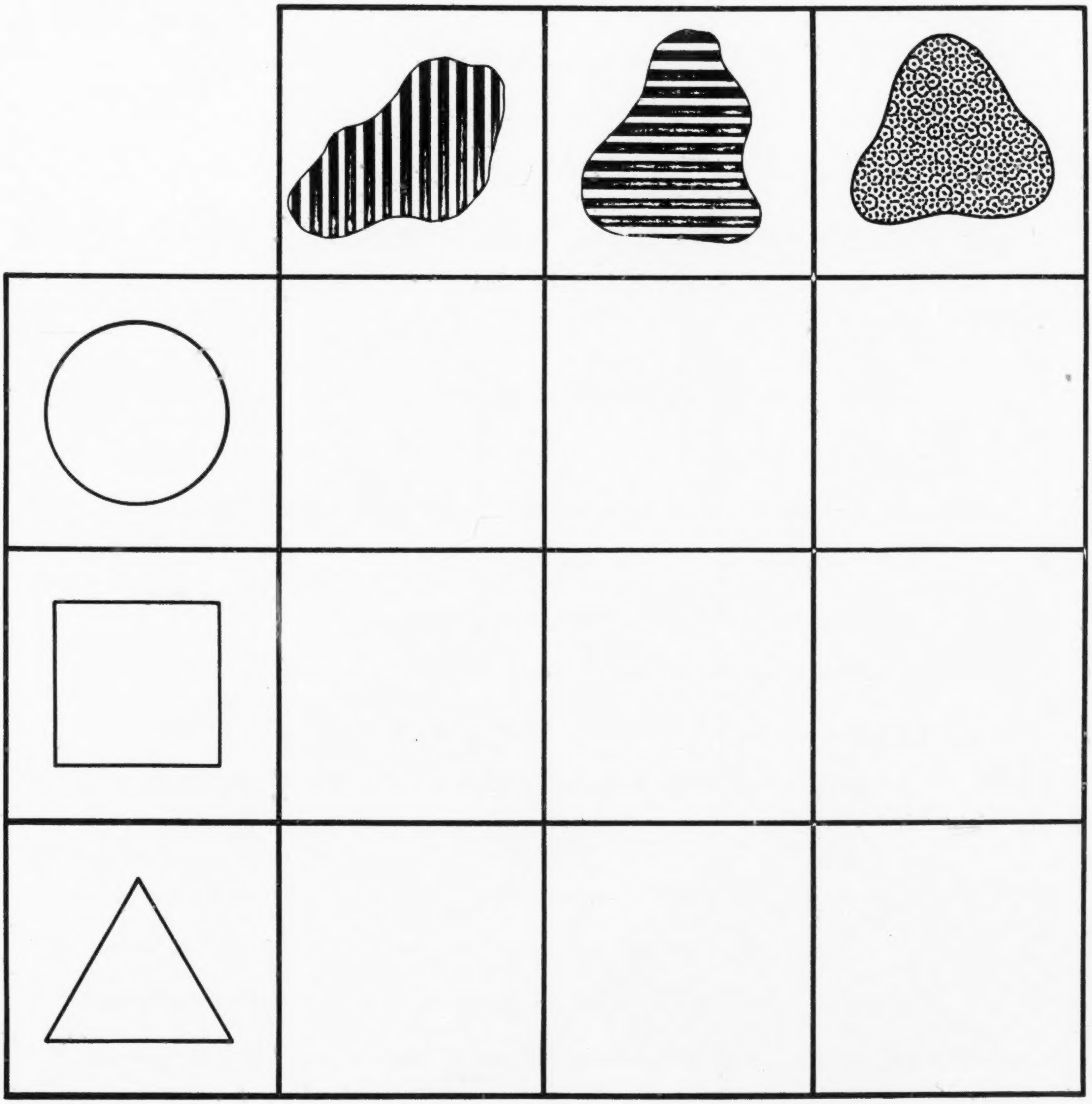


Figure 1: Stimulus Layout for Attribute Cell Task.

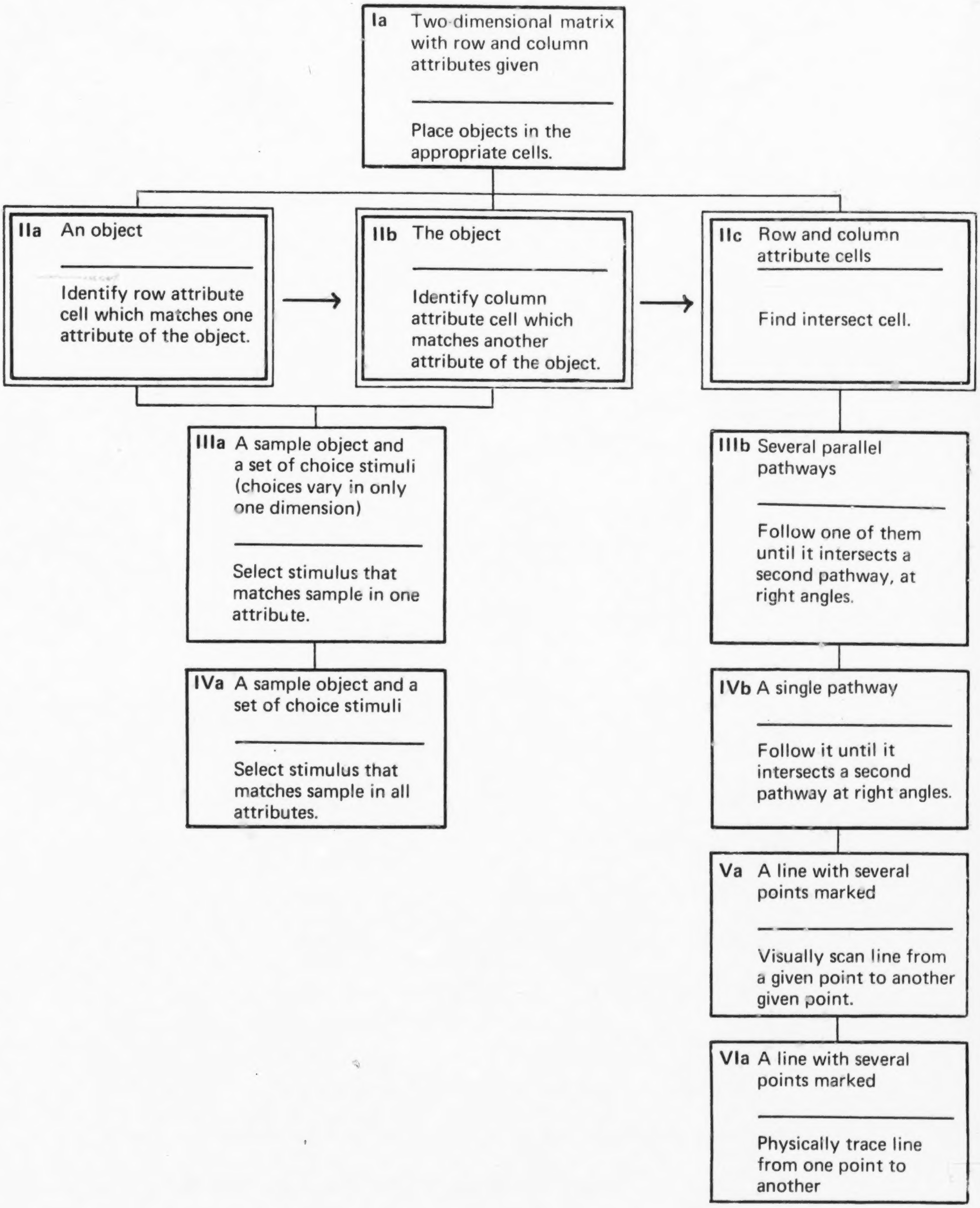

Ia Two-dimensional matrix with row and column attributes given
Place objects in the appropriate cells.
IIa An object
Identify row attribute cell which matches one attribute of the object.
IIb The object
Identify column attribute cell which matches another attribute of the object.
IIc Row and column attribute cells
Find intersect cell.
IIIa A sample object and a set of choice stimuli (choices vary in only one dimension)
Select stimulus that matches sample in one attribute.
IVa A sample object and a set of choice stimuli
Select stimulus that matches sample in all attributes.
IIIb Several parallel pathways
Follow one of them until it intersects a second pathway, at right angles.
IVb A single pathway
Follow it until it intersects a second pathway at right angles.
Va A line with several points marked
Visually scan line from a given point to another given point.
VIa A line with several points marked
Physically trace line from one point to another

Figure 2: Behavioral Analysis for Attribute Cell Task.

Figure 3 represents the stimulus layout for a task in which the child must infer the common row and column attributes on the basis of the arrangement of objects in a partially completed matrix. Behavior analysis of this task, presented in Figure 4, indicates four component behaviors (Boxes IIa-IId). Instead of matching-to-sample, the S must determine what attributes a set of objects has in common (Boxes IIa and b). Hypothesized prerequisites for this behavior are both spatial (IIIa and its prerequisites) and conceptual (IIIb and its prerequisites). An important set of prerequisites involves naming attributes of objects (IVc, Vb, VIa). Thus, some form of linguistic encoding seems necessary to solution, although it should be noted that an S might use "private" rather than standard language labels for the attributes and still solve the matrix task.

Having identified the row and column attributes, the S must next combine the attribute names into a description of an object (Box IIc) and then select the object that meets the description as the appropriate one for the cell (IId). Hypothesized prerequisites for composing the description involve grammatical behavior (Box IIIc), while selecting the appropriate object shares with earlier components in the chain the prerequisites of responding to a verbal label (Boxes Vb and VIa). Thus, these components, too, are heavily linguistic in nature.

These analyses suggest that the Incomplete Matrix Task should be considerably more difficult to learn than the task in which attribute cells are given. However, since the two tasks are similar in stimulus format and logical

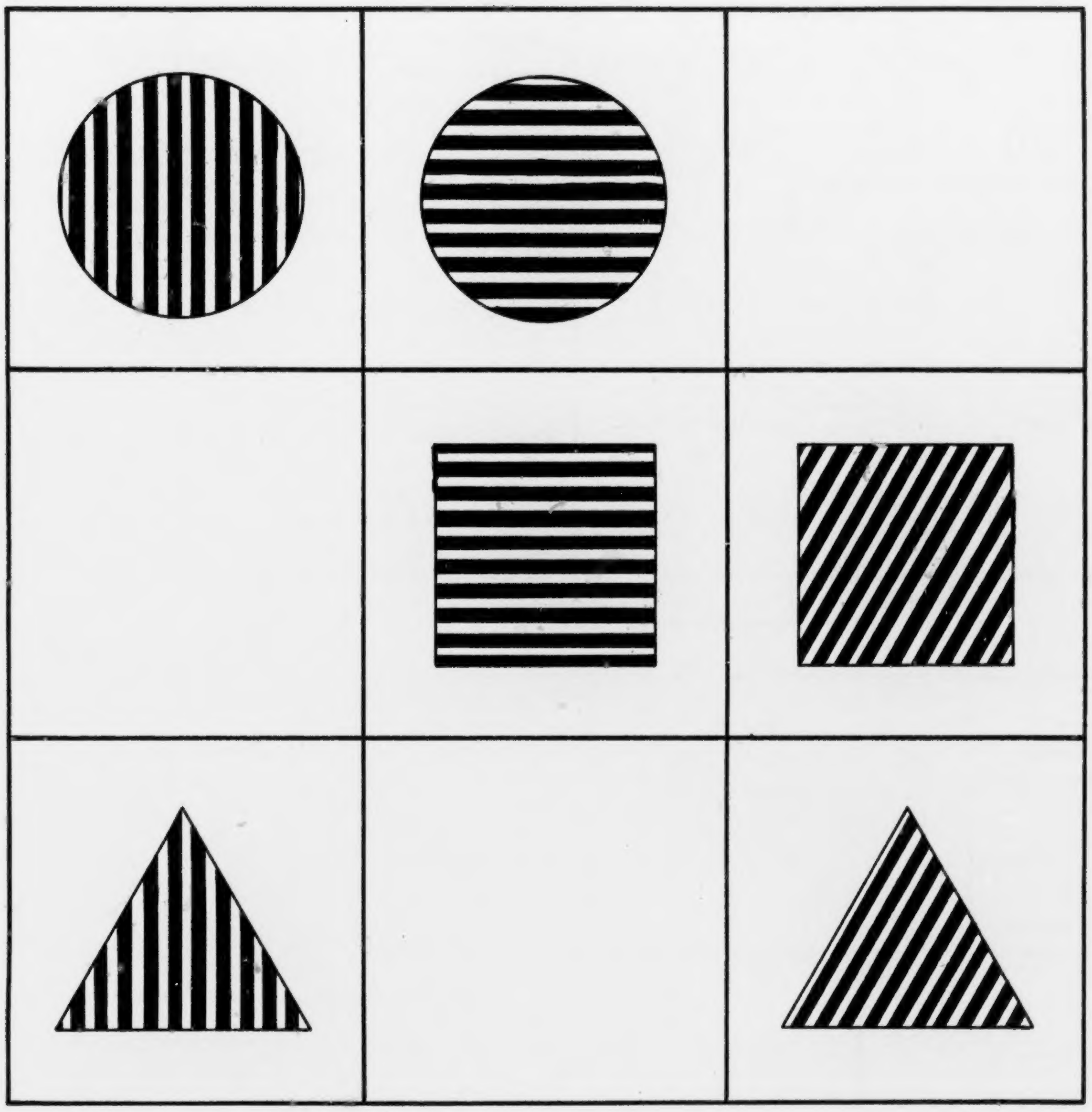


Figure 3: Stimulus Layout for Incomplete Matrix Task.

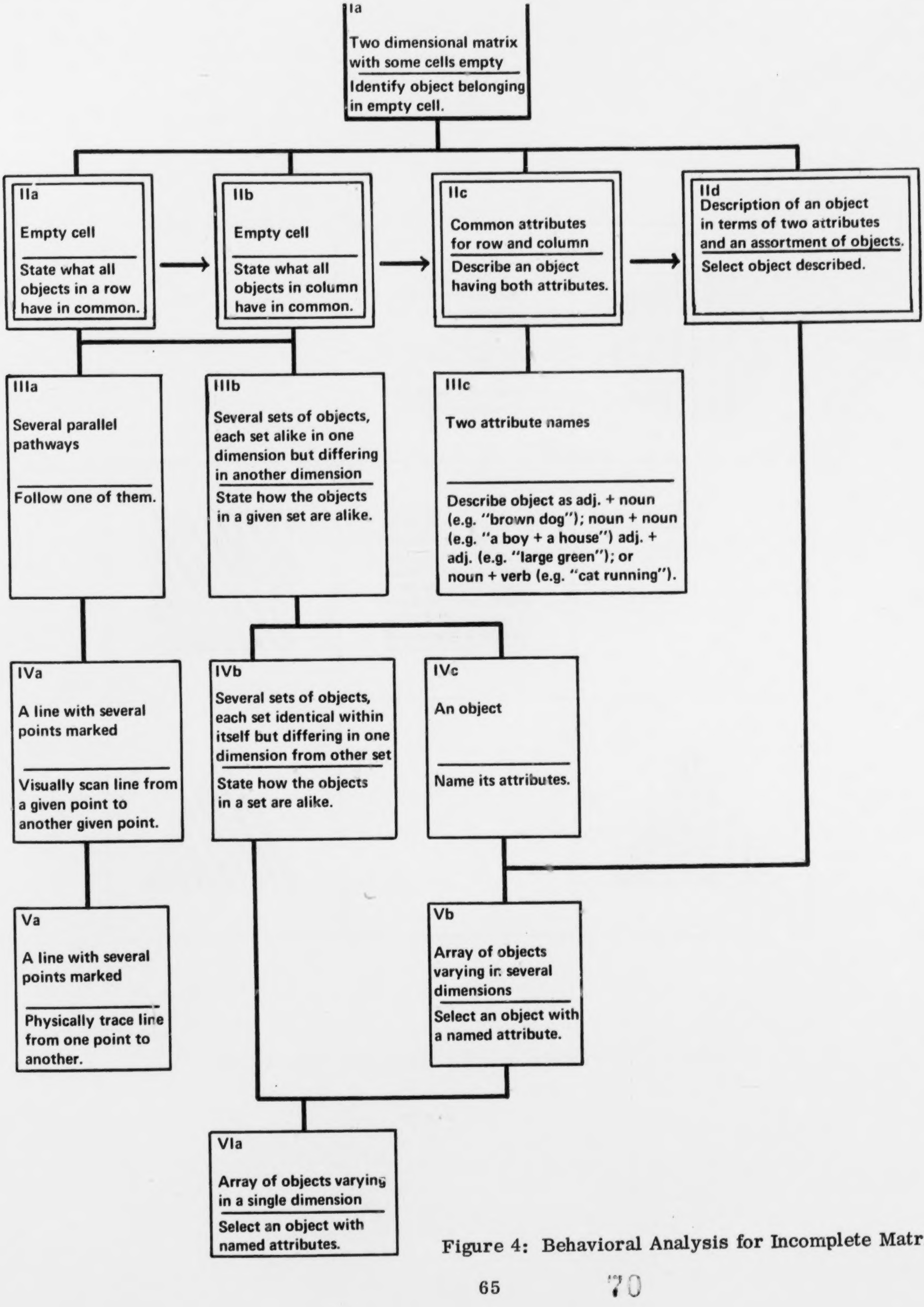

Ia
Two dimensional matrix with some cells empty
Identify object belonging in empty cell.
IIa
Empty cell
State what all objects in a row have in common.
IIb
Empty cell
State what all objects in column have in common.
IIc
Common attributes for row and column
Describe an object having both attributes.
IId
Description of an object in terms of two attributes and an assortment of objects.
Select object described.
IIIa
Several parallel pathways
Follow one of them.
IIIb
Several sets of objects, each set alike in one dimension but differing in another dimension
State how the objects in a given set are alike.
IIIc
Two attribute names
Describe object as adj. + noun (e.g. "brown dog"); noun + noun (e.g. "a boy + a house") adj. + adj. (e.g. "large green"); or noun + verb (e.g. "cat running").
IVa
A line with several points marked
Visually scan line from a given point to another given point.
IVb
Several sets of objects, each set identical within itself but differing in one dimension from other set
State how the objects in a set are alike.
IVc
An object
Name its attributes.
Va
A line with several points marked
Physically trace line from one point to another.
Vb
Array of objects varying in several dimensions
Select an object with a named attribute.
VIa
Array of objects varying in a single dimension
Select an object with named attributes.

Figure 4: Behavioral Analysis for Incomplete Matrix Tas

structure, and since they share the same spatial organization prerequisites, it seems reasonable to assume that learning the easier task first would significantly facilitate learning the Incomplete Matrix Task. The two tasks were, therefore, hypothesized to be hierarchically related, with the Attribute Cell Task prerequisite to the Incomplete Matrix Task. From this general hierarchical hypothesis, three specific hypotheses were derived:

a) The Incomplete Matrix Task will be learned in fewer trials when the Attribute Cell Task has been learned first.

b) Trials to criterion for the two tasks combined will be lower if the tasks are learned in the optimal order (Attribute Cell, then Incomplete Matrix Task) than if they are learned in the reverse order.

c) If the Incomplete Matrix Task is taught first (i.e., non-optimal order), children who succeed in learning it will show nearly immediate mastery of the Attribute Cell Task, since they would have already acquired the elements of the simpler task.

Method

Subjects

Subjects were kindergarten children in a predominantly white, middle-class school. Although 53 children were pretested on the two matrix tasks, only children who failed both were included in the experimental sample. The final sample of 11 boys and 16 girls (ranging in age from 5 years, 3 months to 6 years, 5 months) was matched on total errors on the pretest and then one member of each pair was randomly assigned to each of two treatment groups.

Materials and Task Description

All stimuli and matrices were painted on or cut out of cardboard. The stimuli varied in color and shape.

In addition to the two tasks analyzed above, a form of the Attribute Cell Task with high feedback was used as a Warmup Task. In this task, the child was presented with a three-by-three matrix with filled attribute cells and covered inner cells. E showed the child how the matrix was constructed and named all the colors and shapes. E then told the child that in each box there was an object that had a color and a shape, and that the child's job was to tell him what color and shape it was. After stating his answer for each cell, the child was allowed to lift a flap covering the cell; the correct object appeared underneath.

Six of the nine cells in each matrix were pointed to by E. As was the case with both subsequent tasks, if the child responded correctly, E went on to the next item; if the child responded incorrectly, then a correction procedure was used. The six responses for each matrix constituted one "trial." Criterion was two consecutive successful trials, i.e., two matrices with no more than one error. For this and the subsequent tasks, a maximum of twelve trials was allowed.

In the Attribute Cell Task (Figure 1), the child was presented with a three-by-three matrix in which the attribute cells were filled but the interior cells were empty. As in the Warmup Task, E explained how the matrix was constructed. He then told the child that he would give him an object, and that he (the child) was to put it in the correct cell. E then presented six of the nine objects in random order for each matrix. Each matrix (six responses) constituted a "trial." Criterion was two consecutive successful trials, i.e., two matrices with no more than one error.

In the Incomplete Matrix Task, the child was given a partially filled three-by-three matrix without attribute cells. One, two, or three cells of the matrix were empty. The child was asked to find each missing object from a random array of the nine possible objects for that matrix. Each series of three matrices (total of six responses) constituted a trial. Criterion was two consecutive successful trials; in this case, two series of three matrices with no more than one error.

Procedure

The experimental sessions began ten days after the end of the pretest sessions. Each child was given one or two training sessions per week. Only one task was taught in each session. Once a child reached criterion, the session was ended. If a child did not reach criterion on a task in one session, he was given up to six trials on that task in the next session. After twelve trials or a maximum of four sessions on a task, training was begun on the next task.

There were two experimental groups, defined by the order in which the matrix tasks were taught. Both experimental groups learned the Warm-up Task first. Group A then learned the two experimental tasks in the hypothesized optimal order: Attribute Cell Task first; Incomplete Matrix Task second. Group B learned the tasks in the reverse order. Dependent measures were trials to criterion on each of these two tasks separately, and for the two tasks combined.

Results

The percentages of all Ss (N=53) passing each pretest were: Warmup Task, 36 percent; Attribute Cell Task, 19 percent; and Incomplete Matrix Task, 21 percent. Apparently these tasks presented real challenges for most of the children--only slightly more than a third of the children passed even the easiest task. Only Ss who failed all three pretests were included in the experimental sample (14 in Group A, and 13 assigned to Group B).

Table 1 presents mean trials to criterion on the Warmup Task and the two experimental tasks for both groups of experimental Ss. The difference between the experimental groups in the Warmup Task was not significant ($t < 1.00$), indicating that the groups were equivalent in ability to learn tasks of this type. All but one experimental subject learned this task.

The hypothesis that the group learning the Attribute Cell Task first (Group A) would learn the Incomplete Matrix Task more quickly than the group that began with the Incomplete Matrix Task (Group B) was supported, but not strongly when the data for all 27 children were considered. The acquisition functions for Groups A and B (in terms of cumulative percent of Ss at criterion) are presented in the two solid lines of Figure 5. As is evident in the graph, more Ss reached criterion, and they reached it faster, in Group A than in Group B. However, the difference between the two groups in trials to criterion was only marginally significant ($t=1.36$, df=25, $.05<p<.10$, one-tailed).

TABLE 1

Mean Trials to Criterion on Three Training Tasks

Group	N	Order of Training	Tasks							
			Pretraining		Attribute Cell		Incomplete Matrix		Attribute Cell + Incomplete Matrix	
			$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
A	14	Attribute Cell-Incomplete Matrix	3.43	3.13	4.36	3.58	6.86	3.66	11.21	6.68
B	13	Incomplete Matrix-Attribute Cell	3.85	2.88	5.38	4.16	8.85	3.92	14.23	7.24

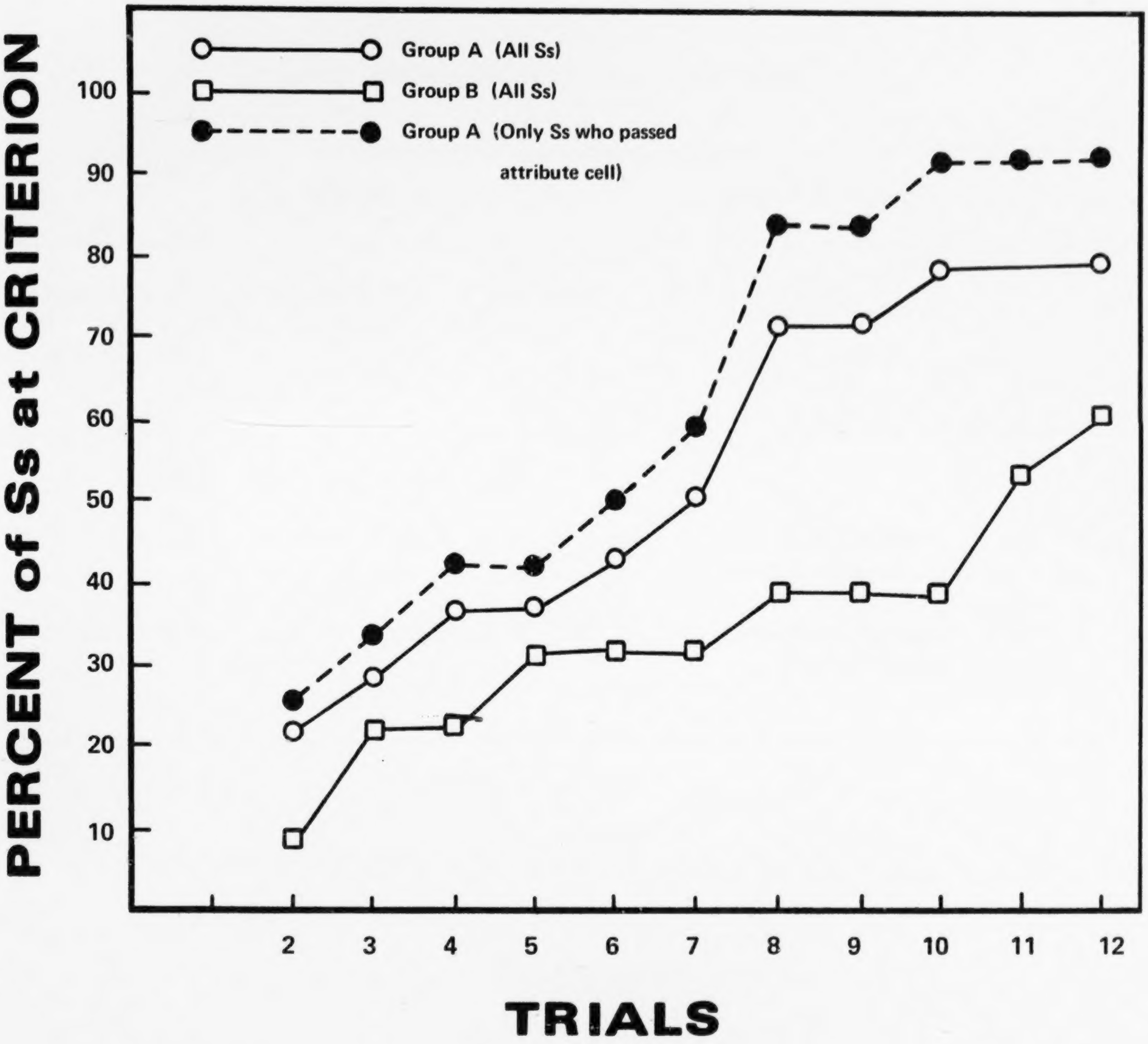


Figure 5: Cumulative Percent Ss at Criterion on Incomplete Matrix Task.

Predictions for this study were specifically concerned with the transfer effects of mastering (as opposed to simply being exposed to) one task on the learning of the next task. Thus, a rigorous test of the hypothesis requires examining, for each successive task, only those Ss who had succeeded in learning the preceding task. This method of analysis treats the data as if any S who failed to learn a task had been dropped from the study and not allowed to proceed to the next task.

The mean number of trials to criterion for the Attribute Cell Task and the Incomplete Matrix Task considering only those Ss who reached criterion on the preceding task appears in Table 2. The dotted line in Figure 5 shows the Incomplete Matrix acquisition function for those Group A Ss who had previously reached criterion on the Attribute Cell Task. As is apparent in the graph, when Ss who failed to learn the hypothesized prerequisite are eliminated from consideration, the difference between Groups A and B becomes much more clearcut. The difference between these Group A Ss and Group B Ss on trials to criterion on the Incomplete Matrix Task was significant at less than the .05 level ($t = 1.92$, df = 22, one-tailed).

A second hypothesis, that Ss who learned the tasks in the optimal order (Group A) would take fewer total trials to learn both tasks than would Ss who learned in reverse order (Group B), was not supported, even though the results were in the predicted direction (11.21 and 12.23 trials, respectively; $t = 1.04$, $p > .10$).

77

TABLE 2

Mean Trials to Criterion for Two Experimental Tasks Considering Only *Ss* Who Reached Criterion on the Preceding Task

Group	Attribute Cell			Incomplete Matrix		
	N	$\bar{X}$	S. D.	*N*	$\bar{X}$	S. D.
A	13	4.23	3.68	11	6.00	3.38
B	8	3.63	3.31	13	8.85	3.92

Our third hypothesis was that the child who learned the Incomplete Matrix Task first would demonstrate almost immediate performance of the easier Attribute Cell Task. A test of this hypothesis requires examining the data for those Ss in Group B who actually learned the Incomplete Matrix Task. Figure 6 represents a data plot for Ss in Group B, with trials to criterion on the Attribute Cell Task on the vertical axis, and trials to criterion on the Incomplete Matrix Task on the horizontal axis. Of the eight children who did learn the Incomplete Matrix Task, six of them took the minimum possible number of trials (two) to learn the Attribute Cell Task--that is, they learned the easier task immediately.

Discussion

These results generally confirm the hypothesized hierarchical relationship between the two matrix tasks, and thus lend support to the technique of behavior analysis used to generate the hierarchy. Although the two experimental tasks seem to be very similar superficially, behavior analysis suggested that the Incomplete Matrix Task required all the critical components of the Attribute Cell Task plus the additional linguistic components of discovering and naming common attribute values, and composing verbal descriptions of the objects. Thus, the Incomplete Matrix Task was placed above the Attribute Cell Task in the hierarchy, implying that prior learning of the Attribute Cell Task would facilitate learning the Incomplete Matrix Task. The advantage of Group A over Group B in learning the Incomplete Matrix Task confirms this. Further support comes from the fact that six out of eight children who first mastered the more difficult task made no errors on the easier task.

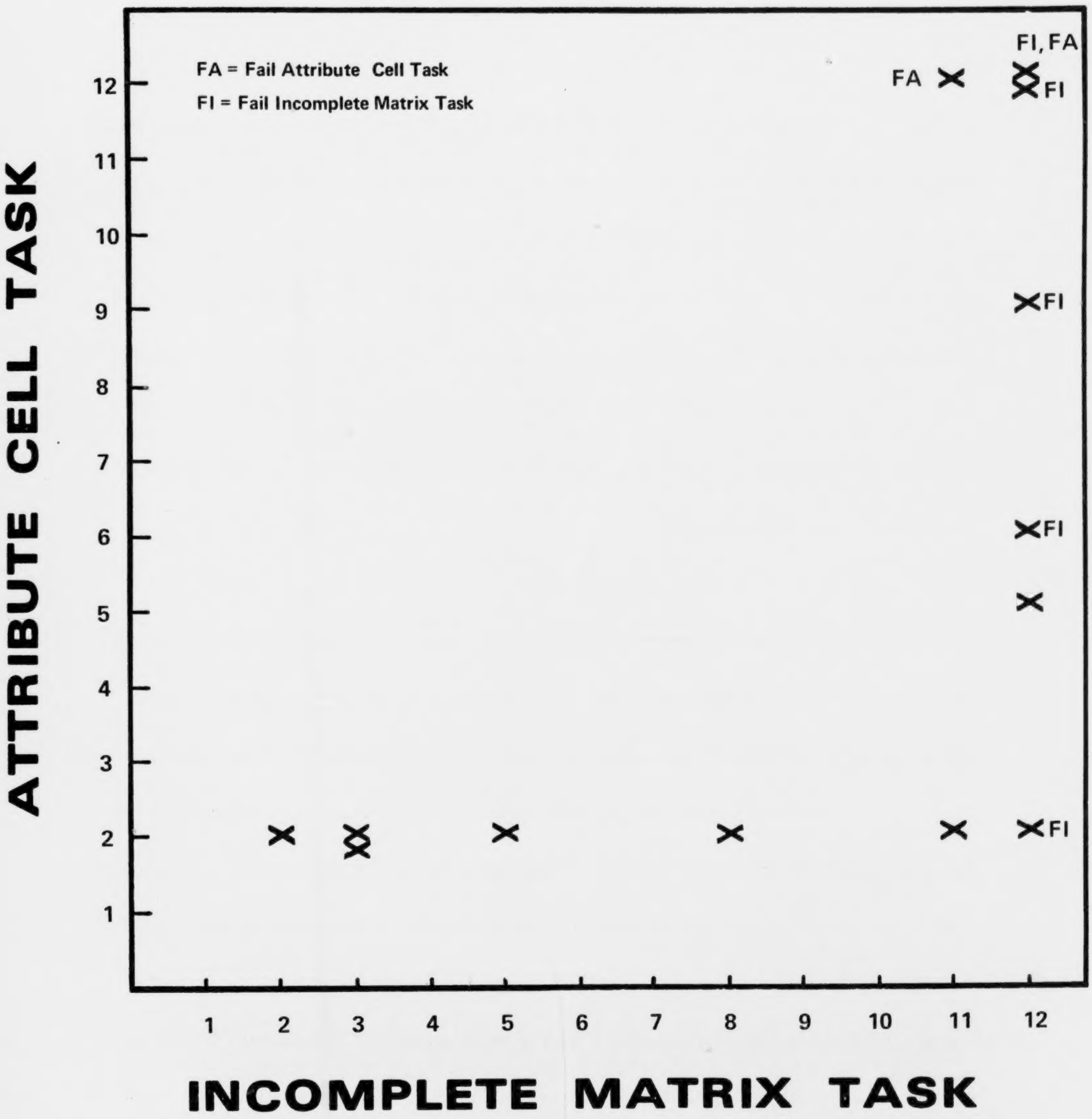


Figure 6: Trials to Criterion on Attribute Cell and Incomplete Matrix Tasks for Group B.

This study also demonstrates that "pre-operational" children can learn "concrete-operational" skills when they are given the opportunity to learn component and prerequisite behaviors through corrected practice on a series of simpler, related tasks. This is in accord with studies of "programming" and successive approximation in children's learning of discriminations, and suggests that acquisition of more complex cognitive skills, as well as simple discriminations, may be a matter of learning specific prerequisites, rather than of entering a general level or "stage" of development. At the present, however, the burden of proof for this contention remains the authors'.

References

Bruner, J. S., & Kenney, H. J. On multiple ordering. In J. S. Bruner, R. R. Olver, & P. M. Greenfield (Eds.), Studies in cognitive growth. New York: Wiley, 1966.

Flavell, J. H. The developmental psychology of Jean Piaget. New York: D. Van Nostrand, 1963.

Gagné, R. M. The acquisition of knowledge. Psychological Review, 1962, 69, 355-365.

Gagné, R. M., Mayor, J. R., Garstens, H. L., & Paradise, N. E. Factors in acquiring knowledge of a mathematical task. Psychological Monographs, 1962, 76 (Whole No. 526).

Gagné, R. M., & Paradise, N. E. Abilities and learning sets in knowledge acquisition. Psychological Monographs, 1961, 75 (Whole No. 518).

Inhelder, B., & Piaget, J. The early growth of logic in the child: Classification and seriation. New York: Harper & Row, 1964.

Lovell, K., Mitchell, B., & Everett, I. R. An experimental study of the growth of some logical structures. British Journal of Psychology, 1962, 53, 175-188.

Overton, W., & Brodzinsky, D. Perceptual and verbal factors in the development of multiplicative classification. Paper presented at meetings of the Eastern Psychological Association, April, 1969.

Resnick, L.B., Siegel, A. W., & Kresh, E. Transfer and sequence in learning double classification skills. Journal of Experimental Child Psychology, 1971, 11, 139-149.

Shantz, C. U. A developmental study of Piaget's theory of logical multiplication. Merrill-Palmer Quarterly, 1967, 13, 121-137.

Siegel, A. W., & Kresh, E. Children's ability to operate within a matrix: A developmental study. Developmental Psychology, 1971, 4 (2), 232-239.

Smedslund, J. Concrete reasoning: A study of intellectual development. Monographs of the Society for Research in Child Development, 1964, 29 (2, Whole No. 93).

Smedslund, J. Determinants of performance on double classification tasks. I. Effects of covered vs. uncovered materials, labeling vs. perceptual matching, and age. Scandinavian Journal of Psychology, 1967, 8, 88-96. (a)

Smedslund, J. Determinants of performance on double classification tasks. II. Effects of direct perception and of words with specific, general and no reference. Scandinavian Journal of Psychology, 1967, 8, 97-101. (b)

A STUDY OF SUBORDINATE SKILLS IN SCIENCE PROBLEM SOLVING[1]

Virginia K. Wiegand

University of California, Berkeley[2]

In this study, the maturational view of Piaget was compared with that of cumulative learning with respect to solving a science problem. The success of instructional programs based on learning hierarchies in enabling students to perform conservation and mathematical tasks (e.g., Gagné, 1962; Gagné & Brown, 1961; Gagné & Paradise, 1961; Gagné, Mayor, Garstens, & Paradise, 1962; Kingsley & Hall, 1967; LeFrançois, 1968) made it seem likely that the same general procedure would be equally effective in the learning of a science task requiring logical thinking. Further, it seemed that the deficiencies in the ability of students to solve this particular type of science problem could best be described as absences of specific capabilities or subordinate intellectual skills, each of which could be overcome by specific learning. The individual assessment of each student's initial capabilities in order to identify those specific capabilities which he lacked, followed by a learning sequence designed to provide him with the specific learning of any needed subordinate skills, was expected to lead to the successful performance of two previously failed science tasks, a "Final Task" for which a learning hierarchy was constructed, and a "Transfer Task," a problem similar to the Final Task and selected from problems used by Piaget in his studies of logical thinking.

---

[1] This paper is based on a dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of California, Berkeley, California, 1969. The author is greatly indebted to Professor Robert M. Gagné, who served as advisor. This research was supported by the United States Department of Health, Education, and Welfare under Contract No. OEC-4-062949-3066, with Professor Gagné.

[2] Now at California State College, Hayward, California.

Method

Tasks

The transfer problem chosen is discussed in the chapter entitled "Hauling Weight on an Inclined Plane" (Inhelder & Piaget, 1958). In this problem a toy wagon, suspended by a cable, is drawn up an inclined plane by counterweights attached to the other end of the cable. The cable passes over a pulley at the end of the plane. The inclination of the plane can be varied. The child's task is to "predict the movements or equilibrium position of the wagon as a function of three variables--the weight it carries, the counterweight suspended by a cable fastened to the wagon, and the inclination of the track (Inhelder & Piaget, 1958, p. 182)." This problem was adapted for the present study a) by having the child make his observation in a predetermined specified order; b) by having him tabulate the results of his observations; and c) by using vertical heights of 4, 5, and 6 inches for the inclination of the plane rather than the inclinations of 30, 45, 60, and 90 degrees commonly used by Piaget's subjects. After completing the series of observations and tabulations, each subject in this study was asked to derive from his results the general equation $H \times W = T \times C$, where H is the vertical height of the incline, W is the weight of the car, T is the length of the track along the incline from the horizontal to the position of equilibrium, and C is the weight of the counterweight. This relation could be expressed either in words or symbols. Figure 1 shows an example of one of the problems posed for this "Transfer Task."

A "Final Task" involving the inclined plane was employed as an immediate measure of the effects of the experimental treatments. In the Final Task, the small car was placed at a designated place on the incline. When the car was released, it traveled down the inclined plane and struck a block placed at the bottom of the incline along the horizontal. The student was asked to derive from his observations and tabulations the relation between a) the vertical height of the incline at the starting position of the car and the weight of the car, and b) the weight of the block at the bottom of the incline and the distance moved by the block after being struck by the moving car. All combinations of three weights and three heights were used. All observations were made according to a predetermined random order. Figure 2 shows an example of a Final Task problem.

A learning hierarchy of subordinate capabilities was constructed for the Final Task, along with problems used to assess the various subskills. The learning hierarchy is given in Figure 3. For each subordinate capability identified, a set of problems was devised to assess the subject's ability to perform the subskill described. The type of problem used for each subordinate capability is described briefly below.

II. Given a set of values for three or more physical variables (e.g., mass, volume, and density), the student was to place them in order in a table, derive from these values the general equation, then demonstrate the derivation of one

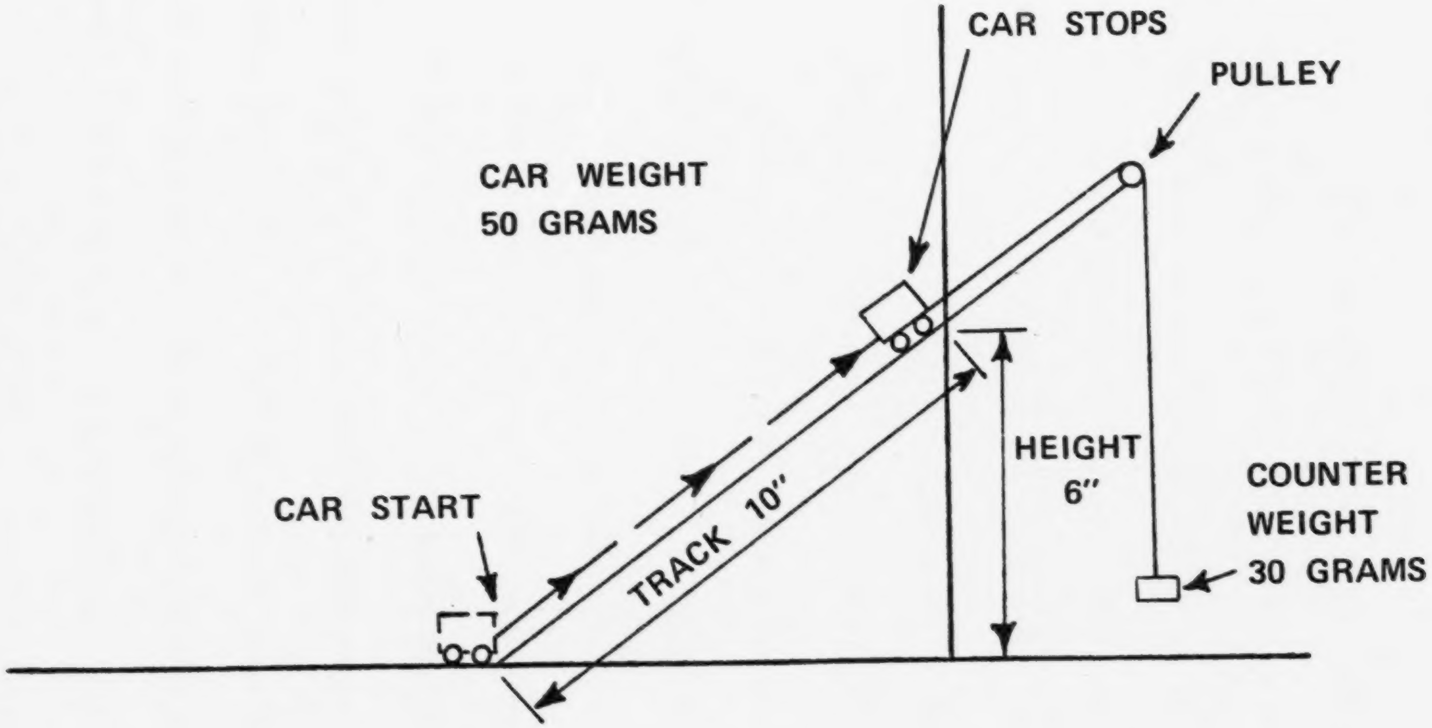

CAR STOPS
PULLEY
CAR WEIGHT
50 GRAMS
HEIGHT
6″
CAR START
TRACK 10″
COUNTER
WEIGHT
30 GRAMS

Figure 1: One of Nine Observations Made in Transfer Task.

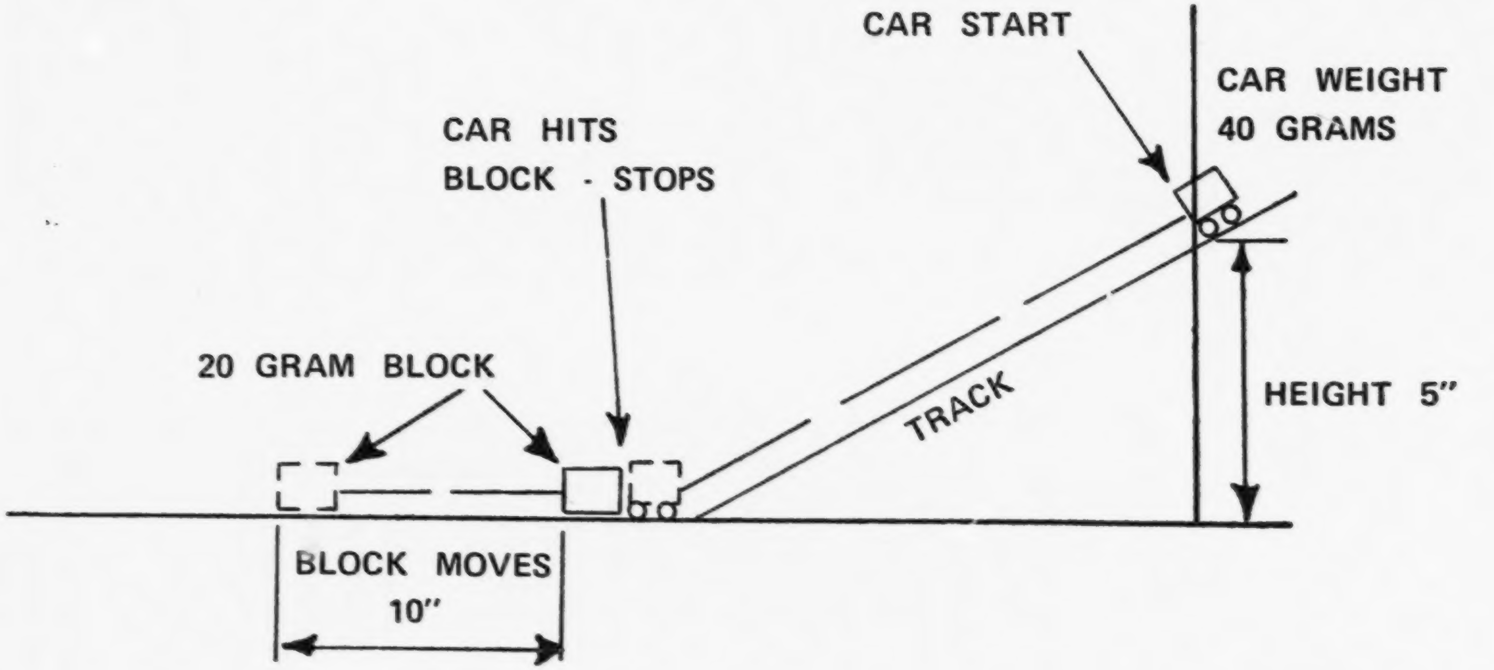

CAR START
CAR WEIGHT
40 GRAMS
CAR HITS
BLOCK - STOPS
20 GRAM BLOCK
TRACK
HEIGHT 5"
BLOCK MOVES
10"

Figure 2; One of Nine Observations Made in Final Task.

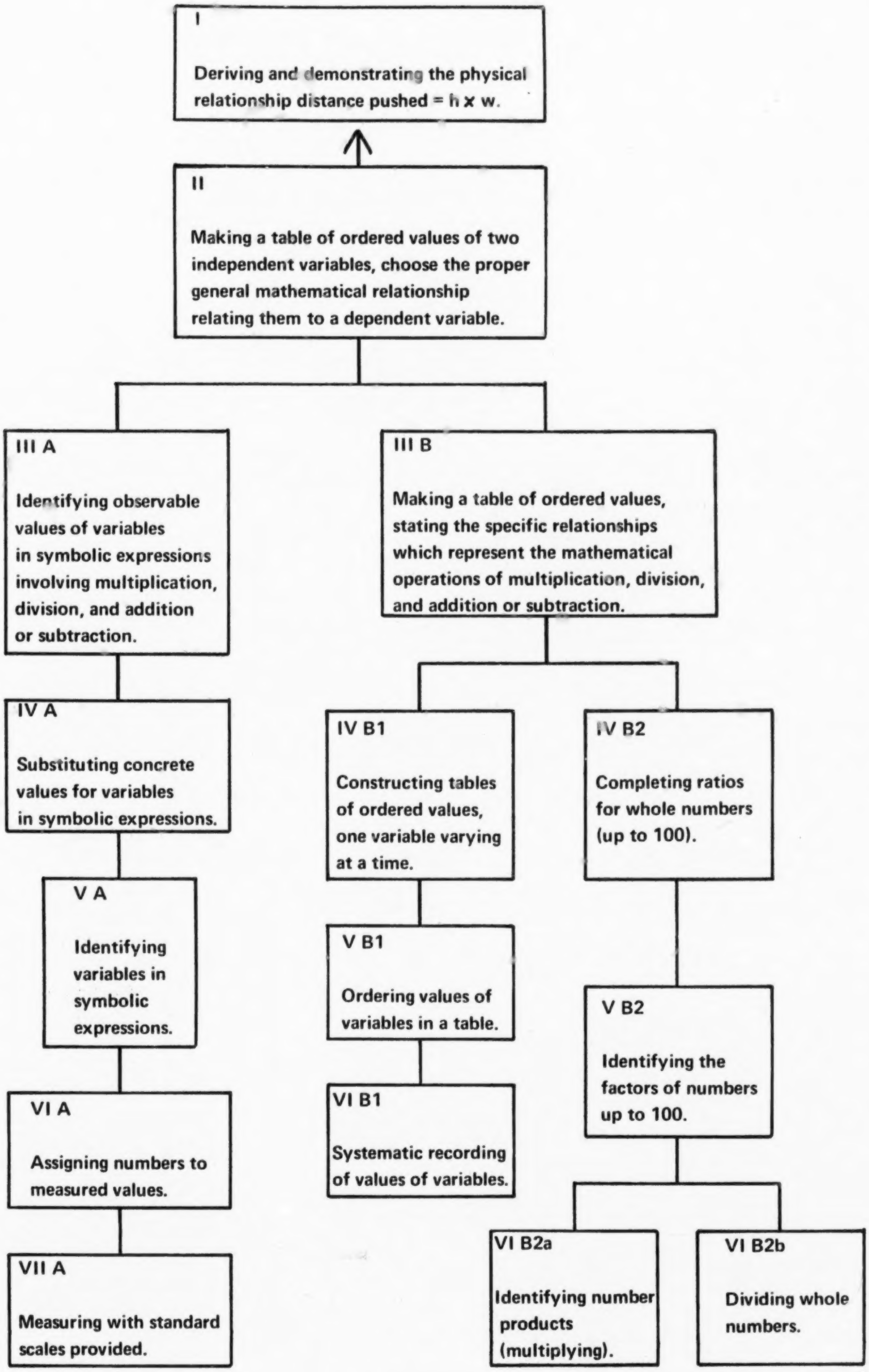

I
Deriving and demonstrating the physical relationship distance pushed = h x w.
II
Making a table of ordered values of two independent variables, choose the proper general mathematical relationship relating them to a dependent variable.
III A
Identifying observable values of variables in symbolic expressions involving multiplication, division, and addition or subtraction.
III B
Making a table of ordered values, stating the specific relationships which represent the mathematical operations of multiplication, division, and addition or subtraction.
IV A
Substituting concrete values for variables in symbolic expressions.
IV B1
Constructing tables of ordered values, one variable varying at a time.
IV B2
Completing ratios for whole numbers (up to 100).
V A
Identifying variables in symbolic expressions.
V B1
Ordering values of variables in a table.
V B2
Identifying the factors of numbers up to 100.
VI A
Assigning numbers to measured values.
VI B1
Systematic recording of values of variables.
VI B2a
Identifying number products (multiplying).
VI B2b
Dividing whole numbers.
VII A
Measuring with standard scales provided.

Figure 3: Learning Hierarchy for a Science Task.

variable given the value of the others. The student had to solve two out of two problems correctly to pass the test of this subskill.

IIIA. Given an equation stated in words (e.g., area = length x width), the student was to show how to find the value of the dependent variable using the equation and values for the independent variables shown in an accompanying illustration. To pass this test, the student had to have two out of three problems correct.

IVA. Given an equation and an assortment of measurements of various types of variables, the student was to substitute the proper numerical values, selected from those given in the assortment of measurements, into the equation. To demonstrate this subskill satisfactorily, the student had to score two of three problems correct.

VA. Given an equation and an illustration, the student labeled the illustration with the proper letters symbolizing designated variables. The criterion for this subskill was two of three problems correct.

VIA. Given two sets of illustrations, one showing an object being weighed and one showing an arrow pointing to a value on a number line, the student was to assign correct numerical values to the weight of the object and the position of the arrow. Two problems, each consisting of five items, had to be done correctly for the student to pass the test of this subskill.

IIIB. Given a set of values relating two physical variables, the student placed them in a table in order and derived from these values the general equation illustrated by the values. The student had to get two of three problems correct to demonstrate this subskill.

IVB1. Given a set of values relating three or more physical variables, but varying values of only one variable, the student was to place them in order in a table and derive from these values the general equation. To pass this test, the student had to get two of three problems correct.

VB1. Given a set of values relating two physical variables, the student was to place them in order in a table for which columns and headings were provided. To pass the test, the student had to be correct for two of three problems.

VIB1. Given a set of values relating two variables, the student was to record the values of the variables in a table for which columns and headings were provided. Two of two problems correct were required for this test.

IVB2. Given three of four values of a ratio, the student was to complete the ratio. To pass the test, four of five problems had to be completed correctly.

VB2. Given a number, the student was to supply the required number of factors (e.g., 32 = __x__x__). Criterion was two of three problems correct.

VIB2a. Given two numbers, the student was to find their product. To pass the test, two of three problems had to be done correctly.

VIB2b. Given a dividend and a divisor, the student was to find the quotient. The criterion was two of three problems correct.

All tests for subskills contained in the learning hierarchy were scored either pass or fail. The Final Task and the Transfer Task were scored in two ways, pass or fail, and in points. For these tasks, one point was given for the correct recording and tabling of values, one point for the correct ordering of the values, and one point for the correct derivation of the general rule; a total score of three points could be attained on either task.

Procedure and Subjects

There were three phases in the experiment: a) pretesting on the Final Task and the Transfer Task, b) assessment of subordinate capabilities, and c) posttesting on the Final and Transfer tasks. Thirty subjects approximately twelve years of age who failed both the Final and the Transfer tasks on a pretest were assigned to one of the three experimental conditions: Demonstration-Test-Retest (DTR), Test-Retest (TR), and Test (T). Pre- and posttest conditions for the Final and Transfer tasks were identical for all groups. Difference in treatment lay entirely in the way in which assessment of subordinate capabilities was accomplished.

In the DTR condition, a demonstration of an example and an explanation of the steps in its solution were included in the instructions given during the initial test of each subordinate capability. The initial test was followed by a retest, without demonstration, of those capabilities failed in the initial test. The initial test began with the most complex task in the learning hierarchy and was continued, moving down the hierarchy, until the student was successful on two consecutive subordinate tasks in each branch of the hierarchy. In the retest, for each branch the simplest task previously failed by the student was given first, and the order of presentation of subsequent problems was the reverse of that in the initial test. The same problems and instructions were used in the initial test and retest.

In the TR condition, the initial test and the retest were given, using the same problem sequence as for DTR. However, there were no demonstrations during the initial test.

In the T condition, only the initial test was given, beginning with the most complex task and moving down the hierarchy.

Results

Examination of the pattern of passing and failing scores on the initial test of subordinate skills showed only three cases (all in group TR) of a total of 172 observations in which a subject passed a lower-level test while failing a related adjacent higher-level test. For the retest scores for Groups DTR and TR there were only two such cases (one in each group). In addition, no subject passed the posttest on the final task without also passing either the initial test or retest for subordinate capabilities II, IIIa, and IIIb. These results show that higher-order skills are indicative of the prior attainment of simpler related skills.

Table 1 shows the number of subjects in each group failing each of the subordinate skills on the initial test and the retest. There was no difference between group DTR and the other two groups on number of initial tests passed ($\chi^2 = 3.84$; df = 1; $p > .05$). This indicates that the demonstration had no significant effect on performance on the initial test.

The retest of subordinate capabilities failed in the initial test enabled subjects to acquire needed subskills, as can readily be seen in Table 1. In groups DTR and TR there were 33 failures in the initial test of subordinate capabilities. There were two failures in the retest for the same groups.

TABLE 1

Number of Subjects Failing Initial Tests and Retests of Subordinate Capabilities

Skill	Group DTR		Group TR		Group T
	Initial test	Retest	Initial test	Retest	Initial test
II	5	0	7	0	7
IIIA	1	0	0	0	1
IVA	0	0	3	0	1
VA	0	0	0	0	0
VIA	0	0	0	0	0
IIIB	3	0	6	0	6
IVB1	2	0	1	0	1
VB1	0	0	0	0	0
VIB1	0	0	0	0	0
IVB2	2	1	2	1	2
VB2	1	0	0	0	0
VIB2a	0	0	0	0	0
VIB2b	0	0	0	0	0
Total	14	1	19	1	18

Table 2 shows the number of subjects in each group who passed the Final Task and the Transfer Task posttests. Nineteen of 20 subjects in Groups DTR and TR were able to pass the Final Task after having attained the needed subskills through the retesting procedure. Subjects successful in performing the Final Task were, in all cases, successful in performing the Transfer Task. Of Group T subjects, only three, each of whom had passed the initial test for subordinate capability II, were able to pass the Final and Transfer Task posttests. This indicates that simply *taking* the subordinate tests did not enable subjects to pass the Final Task. It was necessary to *pass* each subordinate task, either on initial or retest.

Discussion

In this experiment, children who could perform neither a Piaget task (the Transfer Task) nor a very similar science problem (the Final Task) learned, in the space of a few hours, to perform both of these tasks, providing they had also learned any needed subordinate capabilities in the interval between the first and second presentations of the tasks. The experiments also showed that successful performance of subordinate tasks was sufficient in itself to produce transfer to the next higher-level task. There was no direct instruction prior to the retest and, in fact, demonstration prior to the initial test proved of no value to subjects.

Considered in a general sense, the results of the study are consistent with a view of intellectual development that contrasts with that of Piaget. The inability of 30 out of 31 sixth-graders to initially perform the two criterion tasks

TABLE 2

Comparison of Successes on Final and Transfer Tasks with Performance on Retest for Groups DTR and TR and Initial Test for Group T

Group	Subordinate Task II (Initial Test or Retest)	Final Task	Transfer Task
DTR	10	10	10
TR	10	9	9
T	3	3	3
Total	23	22	22

did not appear to be related to a deficiency in logical thought processes. These children were apparently quite capable of engaging in rational thinking. What was lacking in varying degrees, however, was a set of specific intellectual skills pertaining to the use of ratios, the systematic tabulation of data, and the formulation of simple equations. When these subordinate skills were established by means of step-by-step assessment procedures, possibly bringing about recall in some instances and learning in others, it became possible for the children to solve the criterion problems with great competence. These results indicate that intellectual development has been brought about by the cumulative effects of the learning of concretely referenced intellectual skills, rather than by the adaptation of structures of intellectual growth.

References

Gagné, R. M. The acquisition of knowledge. Psychological Review, 1962, 69, 355-365.

Gagné, R. M., & Brown, L. T. Some factors in the programming of conceptual learning. Journal of Experimental Psychology, 1961, 62, 313-321.

Gagné, R. M., Mayor, J. R., Garstens, H. L., & Paradise, N. E. Factors in acquiring knowledge of a mathematical task. Psychological Monographs, 1962, 76 (Whole No. 526).

Gagné, R. M., & Paradise, N. E. Abilities and learning sets in knowledge acquisition. Psychological Monographs, 1961, 75 (Whole No. 518).

Inhelder, B., & Piaget, J. The growth of logical thinking from childhood to adolescence. New York: Basic Books, 1958.

Kingsley, R. C., & Hall, V. C. Training conservation through the use of learning sets. Child Development, 1967, 38, 1111-1126.

LeFrançois, G. A treatment hierarchy for the acceleration of conservation of substance. Canadian Journal of Psychology, 1968, 22, 277-284.

DISCUSSION

John B. Carroll

Educational Testing Service

Let me focus on defining a learning hierarchy. Lauren Resnick distinguishes several possible meanings for a hierarchical relation between two tasks: a) the transfer definition, b) the psychometric definition, and c) the developmental definition. I would say that of these three, (a) is the only one that really stands up, and it stands up because it incorporates a way of testing whether ability in one task is really prerequisite to performance in another task. In fact, I would prefer to think of a learning hierarchy in terms of a network of relations of prerequisiteness. I find that I have to define prerequisiteness in a special way, as will be seen when we look at the developmental definition.

The transfer definition says that two tasks are hierarchically related when learning one task produces positive transfer to the other; one assumes incidentally that the experiment demonstrating the transfer has proper controls, such that individuals not having competence in the "easier" task will reliably fail the "more difficult" task. I should remark also that the relative ease or difficulty of the tasks is only an incidental aspect of the hierarchical relation between them, and that how we define ease or difficulty makes a difference. If relative ease or difficulty is defined in psychometric terms, i.e., relative number of people able to perform the items, it is merely a logical consequence of the prerequisite relation that fewer out of a given sample can perform the "harder" task. But if we define ease or difficulty in terms of "ease of learning," either one of the tasks might be "easier" in this sense. One can conceive, for example, that the task that is lower in the hierarchy is the more "difficult to learn"--while the higher task is quite "easy to learn" for individuals who have learned the task lower in the hierarchy. For example, if "learning the rules of a game" is the lower-order task, and "playing the game" is the higher-order task, it might be that the rules are quite difficult to learn, but playing the game would be quite easy once the rules are learned.

We might stop at this point to inquire further what a prerequisiteness relation between two tasks would mean. If transfer between the two tasks can be demonstrated, it would mean that the lower-order task is a component of the higher-order task--and that the higher-order task is more complex than the other. In fact, the higher-order task might be so complex that a large number of components would be involved. These components might or might not be hierarchically related among themselves. Driving an automobile is a good example of a complex task, with a number of components that are largely independent of one another--being able to start the car, being able to steer accurately, knowing the rules of the road, etc. In the cognitive field, being able to speak a foreign language would be an example of such a complex task--dependent upon knowledge of various aspects of the language system, etc.

Now let us examine the psychometric definition of a hierarchical relation. Again, the relative "psychometric" ease or difficulty of the two tasks is incidental and only a logical consequence of the basic definition: that all who can perform one task (the higher-order one) will also be able to perform the other (lower-order) task, and (I would add) all who fail the lower-order task will also fail the higher-order task. Many pairs of tasks can be found that have this relation, but not all of them would necessarily have a transfer relationship. A simple case of one that does not would be that in which the same task is defined with two criteria of performance--e.g., the running high-jump first at four feet, and then at six feet. All who can jump at six feet can certainly jump at four feet, and those who cannot jump at four feet can certainly not jump at six feet. But I doubt that learning to jump at four feet will necessarily transfer (in the strict sense) to learning to jump at six feet. That is, in learning to jump at six feet, people who have not been trained to jump at four feet might learn to jump at six feet just as readily as those who have been trained to jump at four feet. Jumping at four feet is not necessarily a "component" of jumping at six feet. In any case, it seems to me that we are not much interested in what I would call "natural hierarchies" defined in terms of different criteria (speed, facility, etc.) of performance for the same basic task.

A better case against the psychometric definition can be made when we consider two tasks that merely represent different levels of education or experience. If Task A is "adding two one-digit numbers" and Task B is "passing a high-level vocabulary test," and if we take a broad population of children and adults, we might find the psychometric definition of a hierarchical relation between the two tasks to hold, even though the transfer definition would hardly be reasonable. Being able to add numbers is not logically a component of vocabulary knowledge. Thus, I feel that the psychometric definition is only useful as a heuristic device, i.e., in searching for pairs of tasks that might be tested for hierarchical relations by the transfer criterion. The psychometric definition is only useful because any tasks that are hierarchical by the transfer criterion would also be hierarchical by a psychometric criterion.

Now let us look at the developmental definition, which says that two kinds of performance are hierarchically related when being able to perform the first one precedes the other one in a developmental or maturational sense. In my opinion, the use of the idea of "hierarchy" is inappropriate in this context, if "hierarchy" implies a component relation. Those who work with Piagetian stages would recognize that frequently, the kind of behavior observed in the early stage is completely lost when the later stage is reached. In the conservation experiment, the child who has reached the concrete operation stage will not exhibit the kind of responses he exhibited in the pre-operational stage--the behavior is qualitatively different. He no longer asserts that there is a different amount of water when it is poured from a low wide jar to a tall narrow one. To be sure, passing through the "pre-operational" stage may be "prerequisite" to the "concrete operational" stage, but only because the stages are ordered developmentally. Thus, we must rule out "prerequisiteness" in the developmental sense. There would be no way of testing transfer, because there is no way of arranging that a child not pass through a given stage. Similar remarks could be made about the stages of child language acquisition. This is not to say, of course, that information about developmental stages would be of no use in constructing curricula.

On the basis of the above, I'm pleased to note the papers of the present symposium that emphasize the transfer definition of learning hierarchies. It is of course unfortunate that learning hierarchies cannot be defined and validated except by the transfer experiment, because transfer experiments are difficult and time-consuming. As I said, the psychometric definition is useful as a heuristic device. In that connection, I suggest that multi-dimensional scaling techniques be used to test non-hierarchies as well as hierarchies. For example, in the data presented by Margaret Wang, I would trust that the MSA technique would reveal a nonreproducible scale if one chooses a set of tasks that are not on a hypothesized hierarchy.

I have done some work on further psychometric approaches to hierarchies, but in view of my above remarks I'm not too happy with them. What I investigated in particular was the possibility of identifying disjunctive hierarchical relations on a psychometric basis--cases in which either of two tasks A and B (but not necessarily both) might be presumed prerequisite to some more complex task C. I was unable to find any such disjunctive relations in Gagné's data on mathematical tasks, and the ones I found in some of Margaret Wang's data were difficult to interpret, possibly because of unreliability in the data and the small number of cases. Thinking about this further, I believe it unlikely that valid disjunctive hierarchies will be found: if they are, it might mean that the higher-order task is actually different, depending upon what prerequisites enter into it. In other words, Task C is only phenotypically the same as Task C' if A is prerequisite to C and B is prerequisite to Task C', where C and C' are initially presumed to be the same task. The more likely case is the conjunctive hierarchy, where *both* A and B are prerequisite to C. I am planning to investigate this further, using psychometric procedures. But even if I find such conjunctive relations psychometrically, they would need to be validated by a transfer experiment in which training on both tasks A and B (as opposed to only one) would be tested for transfer to Task C.

A couple of comments on other papers: In the Wiegand experiment, it seems to me that what was demonstrated was not that *learning* of the component skills was prerequisite to the "final" and "transfer" tasks, but that immediate *experience* with the component skills was prerequisite. The data show that most of the pupils could perform most of the component skills anyway, particularly if they had the opportunity at least to practice each skill twice (with or without a demonstration). The demonstration and testing of the subordinate skills helped the pupils to recognize their applicability in the criterion tasks, and thus to follow an analytical method in solving the criterion tasks that they were not able to follow when first confronted with them. A similar situation may have obtained in the Siegel and Kresh experiment--experience in the "placing" phase helps the child analyze the "inferring" phase. What I am suggesting is that the "subordinate skills" as such may not be as relevant as some more general competence in analysis of a complex task that is somehow gained by experience with the subordinate tasks. I am sure the very competent investigators who have presented papers today can think of ways to test this hypothesis, which puts a somewhat different interpretation on the learning hierarchies which have been presumed by them.

DISCUSSION

Robert Glaser

University of Pittsburgh

First, let me quibble a little and then go on to some broader issues. In her remarks, Resnick makes the point that a hierarchy of learning tasks need not be linear. This same point is made by Wang and Siegel. They appear to use linear not in the sense of a linear equation; it is used more in the sense of a temporal sequence or a straight line on a chart, so that a hierarchy chart is either a straight line or a tree structure. The tree structure implies essentially that certain subtasks can be acquired independently of each other, but that they still comprise sequences that need to be learned prior to a superordinate behavior. In this sense, nonlinearity means that there are optional learning paths, but still that any one path taken by a learner is a linear one, even if the transfer between two behaviors is so rapid that they appear to be learned at the same time. Nonlinearity in a stricter sense would imply that the learner can arbitrarily move between superordinate and subordinate tasks, which is not what is implied in the learning hierarchies postulated.

Resnick suggests a "rigorous" technique for specifying the component behaviors and subchains in a hierarchy. Siegel refers several times to "systematic behavior analysis" procedures. In their presentations I have not seen suggested a set of rules by which a hierarchical analysis is rigorously and systematically performed, other than perhaps two rules: a) lay out a chain and b) ask the question originally asked by Gagné, "What kind of capability would an individual have to possess to be able to perform this task successfully, were we to give him only instructions?" Resnick notes that this method of hierarchical analysis is limited to behaviors that can be characterized as chains or procedures and is not applicable to "verbal knowledge." Since a child's overt performance is what is described in each of the boxes of a hierarchy and these reduce to procedures, and since the results of verbal knowledge when overtly performed can generally be described as a chain of behavior, I am not sure of the distinction that is being made unless it is between overt and covert performance.

Siegel concludes, as a result of his study, that the acquisition of more complex cognitive skills may be a matter of learning specific relevant prerequisites rather than entering a general level or stage of development. However, does he not need to show that his experiment has not influenced the general level of development of his subjects, or should he not at least assess the level of development that his subjects are in, prior to the experiment?

In Wang's presentation, upon finding an empirical ordering which is different from her postulated ordering, she says that the way in which the task was presented to the subject placed it in a more superordinate category. This is interesting because it indicates that an empirically obtained ordering can sometimes suggest that inappropriate instruction, or inappropriate experimental testing has taken place.

Wiegand concludes in a way similar to Siegel that the results of her experiment show that intellectual development is brought about by the cumulative effects of the learning of concretely referenced intellectual skills rather than by the adaptation of structures of intellectual growth. My own predilections are that this is quite a correct interpretation, but the experimental question remains of the influence of the experiment on whatever can concretely be defined as "structures of intellectual growth." However, I know that it is hard to see the enemies she attacks because they have not defined themselves very concretely or operationally; so I suspect the burden of proof is on them.

Uprichard points out that although the teaching order of "greater than" and "less than" should epistemologically have no effect on learning, his study showed that it was most efficient to learn "greater than" first. This raises the interesting question of the relationship between structures of knowledge as codified and organized for logical, scientific, and theoretical work versus the organization of a body of knowledge for behavioral change, i.e., instructional purposes. Epistemological and psychological structures may not necessarily be the same. Special theories or organizations of knowledge might be designed explicitly for the purpose of teaching a body of knowledge to a novice so that he learns it effectively and gets closer and closer to the expert's organization, and eventually to his own idiosyncratic structure. With respect to this, it is of interest to note that for bodies of knowledge that have been organized by subject-matter specialists, like mathematics, we have a logical structure that can be compared with learning structures. However, for the kind of learning that goes on in very young children, relating to the acquisition of early intellectual skills, we have no guidelines other than developmental theory and experiment, and it is this structure which is fascinating to uncover by attempting to generate and validate learning hierarchies.

I am convinced that learning hierarchy analysis has significant implications for instructional psychology. I suspect that just as the definition of instructional objectives in behaviorally assessable terms has been a powerful factor in the improvement of instructional techniques (frequently more powerful than the manipulation of learning variables commonly studied, like reinforcement delay, spaced practice effects, and so forth), so will the identification of subobjectives and their hierarchical ordering have a powerful influence on the effectiveness of instruction. A special test will be to investigate the relative effectiveness of manipulations of learning variables that are involved in the teaching and learning that fall between the subobjectives as compared with the sole effect of behavioral specification of hierarchical attainments themselves, displayed to the teacher and learner.

Now let me continue with some further considerations on the general topic of learning hierarchies. These are given in the order they occurred to me as I went through the papers. Wang, in her paper, works with two levels of subobjectives which she calls units and subunits. Essentially this means that one hierarchy consisted of finer behaviors within a unit. This brings up the question of the level of analysis one employs in generating a learning hierarchy. It would seem that the grosser the behavioral units employed, the more order one would find, or is this true? Is learning more sequential for fine units than it is for grosser units of behavior? Certainly grosser units of behavior would provide more reliability of measurement, which would keep order measures more consistent; but, on the other hand, the question arises as to whether transfer is more effective among fine than gross units. I am not sure that this is a real question. But, at any rate, I know of no rules that tell us what the appropriate level of analysis is when we generate learning hierarchies. Is it that level that gives us optimal consistency of ordering? Is it that level that gives us information about how to teach the hierarchy? Or, do the units vary as a function of the size of the behavior or the extent of the leap that a learner can take?

The next point involves what can be called "generalized competencies." Not only does a subject learn the behavior specified in the subobjectives, but he also learns some behaviors related to how to go through a learning hierarchy. Two of the papers indicated that some of the subjects appeared to learn "on their own." It seems to me that these "generalized competencies," if they exist, will be confounded with the effects of the cumulative subobjectives the subject learns, so that transfer will be a function of their interaction. This is to say that transfer will be a function of the presence of a subordinate behavior plus the way in which the student acquires the next objective. Resnick has already alluded to the generalized competencies that influence an individual's ability to go through a learning hierarchy in smaller or larger steps, and she has mentioned that there has been little research along these lines. An especially important generalized skill would involve the self-management of one's own learning whereby one generates the size of units and the ordering of units, and sets up a sequence as he learns.

I would like also to refer to a most important and interesting aspect of research on learning hierarchies. This relates to the relationship between the existence of a hierarchy and how one learns the behaviors that comprise it. It can be said that a hierarchy is what exists after learning has taken place. Once a behavior is learned, then the sub-behaviors that comprise it are generally manifested and the presence of a subordinate behavior exists as a component of a superordinate behavior; but, the fact that the behavior is structured in this way may not necessarily mean that it is learned in the way the structure implies or in hierarchical fashion. Learning takes place and then the structure exists; I bumble through a particular set of experiences and after I learn the terminal behavior, the structure is present when someone tests for it.

If what I have just said is true, then a major question follows: given an analyzed hierarchy of behaviors, how does one now generate the most efficient teaching procedure for an individual? The hierarchy gives the subobjective checkpoints, so to speak, but how does learning proceed between or among the checkpoints? Consider some ways in which one might tackle the teaching of a hierarchy. One way is to consider using the powerful technique of observational learning. It is evident that individuals learn from observing others engaged in a performance. The observer learns what behaviors are reinforced, the important discriminative stimuli involved, etc. What effect would a procedure which allowed for observing another child go through a hierarchy have on an individual when he is tested on hierarchy subcomponents? And, what effect would a procedure for a guided and supported engagement in the terminal behavior have?

Another method of instruction might collect all those subobjectives in a hierarchy that comprise the same behavioral class--for example, those that involve making discriminations, those that involve matching-to-sample, those that require learning and using a concept, etc. Supposing one collected the similar classes of behavior that exist in a hierarchy and trained an individual on instances of the class of behavior involved. How would this influence his learning of a hierarchical structure? Another question: What would be the effect of allowing an individual to examine a set of hierarchical behaviors and familiarize himself with them without necessarily attaining mastery of each prior behavior, so that he developed some knowledge of the total structure before beginning to learn the components?

A final implication of the utility of learning hierarchies with respect to test theory and design should be made and Resnick has already alluded to this. A hierarchical structure provides a space in which individuals can be ordered with respect to their level of knowledge in a subject-matter domain.

Each objective in the hierarchy forms a test exercise which an individual can pass or fail; passing means that he should be tested to see what his performance is on the next superordinate objective(s) in order to locate the level at which learning should begin; failing implies that he should be tested on subordinate objectives in order to determine whether his lack of competence is the result of inadequate performance on prerequisite subobjectives, or the result of inadequate instruction on the new objective. Such a structure provides decision paths for the application of sequential or tailored testing procedures where an individual's performance determines his next test exercise. Testing procedures of this kind are practically implemented with a computer and initial research has shown that large savings in testing time can be obtained with reliabilities comparable to tests of greater length. Fundamental to these procedures are necessarily well-worked-out transfer hierarchies.

DISCUSSION

Millie Almy

Teachers College, Columbia University

Dr. Resnick in opening this symposium, and the reports that have followed, demonstrate the parallel interests of learning psychologists, instructional designers, and developmental psychologists. Several of the speakers have suggested that the psychometric and learning studies presented illustrate challenges to cognitive developmental theory as formulated by Piaget or as derived by Kohlberg and others. The potential comprehensiveness of the challenge is apparent in Dr. Wiegand's statement that her results indicate that "intellectual development has been brought about by the cumulative effects of the learning of concretely referenced skills rather than by the adaptation of structures of intellectual growth."

I am a developmental psychologist of Piagetian persuasian and also an early childhood educator with considerable experience in classrooms. I shall maintain this dual stance in considering the very interesting and, I think, extremely useful material presented here.

The cognitive developmentalist is concerned with the child as a knower and with the nature of his knowledge; with the process of acquisition as much as with the outcomes of learning; with changes in the nature of his knowledge over time (over years as well as over days or weeks) and as he reveals it in varied settings.

Cognitive developmental theory is, of course, concerned with learning hierarchies of the sort described here, but as Dr. Resnick's comment on the larger units of analysis implies, the developmentalist places them in a larger context. For Piaget, for example, the fact that a child performs at a particular level in a classification or seriation task is of great significance only if it can be demonstrated that such performance also represents competence--competence that can be spontaneously and appropriately brought to bear on a variety of tasks requiring similar operations in a variety of settings.

Dr. Resnick has analyzed the behaviors involved in a skilled seriation performance and suggests that the errors that characterize the performance of the unskilled might be eliminated through direct instruction at earlier levels. The Piagetian would not deny the efficacy of direct instruction in smoothing particular, specific performances, but would also take note of the contribution of what Kohlberg (1968) calls "general" experience. The skilled seriator, in this view, is one whose environmental encounters, presumably from birth, have been rich and varied. His notion of order is not derived, for example, from mere imitation of the correct procedures for assembling the pink tower or the stairs of the Montessori apparatus, but also involves the information he has received when he has tried to put his own cap on his father's head, or attempted, with difficulty, to lift an object that he has seen his older and stronger brother handle with ease.

The child's competence as a seriator is revealed when he sees possibilities for serial ordering beyond those in which he has received instruction. Take for example the six-year-old who spontaneously takes note of the dates on each of a collection of pennies, arranges them in order, can identify which coin is oldest, which next, and so on, and can specify the number of years between each.

Perhaps it is fair to say that the cognitive developmentalist and the early childhood educator, in contrast to the learning psychologist and instructional designer, are concerned with multiple outcomes. To quote what one educational psychologist has said, "The pupil who moves at a good clip through sequentially organized, mastery-paced arithmetic may be capturing one rule of procedure after another and failing to grasp any intuitive conception of what it is about." He may not be learning to create any procedures for himself. In Piaget's terms, he may not be developing cognitive structures. And furthermore, these procedures and the intimately related affective outcomes cannot be directly taught; they develop by barely perceptible increments. For example, the studies reported by Dr. Siegel and Dr. Uprichard do indeed demonstrate that relatively complex cognitive skills can be learned after specific prerequisite behaviors have been acquired. But neither study provides the information about multiple outcomes the cognitive developmentalist (and the early childhood educator) would like to have. Did the preschoolers go home as conservers, and were the kindergarteners transformed into double classifiers? What, if anything, was the effect on other areas of learning? Suppose the preschoolers, having reached criterion on the set relations tasks, were now confronted with matrix tasks. Would they tackle them as problems involving comparisons? And would the kindergarteners confronted with the set relations tasks start by identifying color and shape?

The point I should like to make here is that children in the early childhood period are extremely good at associative learning, but not very efficient in making appropriate applications in new contexts. Sheldon White (1965) has provided considerable documentation to support the notion that the learning and thinking of the typical eight-year-old differs considerably from that of the typical four- or five-year-old. To put the matter another way, it appears to be relatively easy to program specific behavior sequences for children under the age of seven, and also to demonstrate some degree of transfer, both vertically and horizontally, but I do not think it has been demonstrated that the typical five-year-old can readily be programmed to think like the typical eight-year-old.

Kohlberg's contention that general experience rather than specific learning is essential to move a child to a higher level of thinking seems to me most applicable to the early childhood period. General experience implies the opportunity to be captivated by anything that stirs interest, to explore widely, to make mistakes, and to learn from them. From the viewpoint of the cognitive developmentalist the child's performance at any given point in time is a product of this so-called general experience, and whatever specific instruction he may have had, whether from teachers, parents, other adults, or older children.

In the case of Dr. Uprichard's study one would like to know whether his preschoolers were threes or fours, middle class, or disadvantaged. The different sequences for "equivalence" and "greater than" emerging in his study as compared with Dr. Wang's could reflect differences in the background of experience of the two groups. As Dr. Wang's discussion implies, children who fail a particular pretest may be well advanced along the sequence leading to that task or they may have barely begun it. The process of acquisition may vary accordingly.

Similarly, the success of the testing procedures used in Dr. Wiegand's study may be contingent on the background of experience of the students. As she suggests, certain cognitive structures were already available, needing only concrete support to be applied to science problem solving. Since the age period presumably represented in the sixth grade is regarded as one of attainment for concrete operations and of formation for formal operations, the Piagetian developmentalist would find the results more striking if he knew that the students came from an impoverished environment, or could, as a group, be regarded as slow learners.

Perhaps I have been able to demonstrate that the questions raised by the developmental psychologist and by the early childhood educator are of a somewhat different order from those raised by the learning psychologist and the instructional designer. If this is indeed the case, then hierarchy theory, as it deals with certain aspects of the domain of cognitive developmental psychology, and programmed instruction, as it deals with certain aspects of early childhood education, may be regarded as more complementary than challenging to both those fields.

References

Kohlberg, L. Early education: A cognitive developmental view. *Child Development*, 1968, *39*, 1013-1062.

White, S. H. Evidence for a hierarchical arrangement of learning processes. In L. P. Lipsitt & C. C. Spiker (Eds.), *Advances in child development and behavior*. Vol. 2. New York: Academic Press, 1965. Pp. 187-220.